

INTERNATIONAL APPLICATION PU	BLĮSI	HED (	JN	DER THE PATENT COOPERATION	N TREATY (PC	CT)
(51) International Patent Classification <sup>5</sup> :		·	(1	1) International Publication Number:	WO 94/2	12641
C12N 15/31, C07K 13/00, A61K 39/1 // (C12N 15/31, C12R 1:21)	102	A1	(4	3) International Publication Date:	9 June 1994 (09	9.06.94)
(21) International Application Number: Po	CT/CA	93/0050	01	(81) Designated States: AU, BR, CA, F	I, JP, KR, NO, N	Z, RU,
(22) International Filing Date: 23 November	1993 (2	23.11.9	3)	UA, US, European patent (AT, B GB, GR, IE, IT, LU, MC, NL, P	E, CH, DE, DK, 1	ES, FR
(30) Priority Data: 9224584.4 23 November 1992 (2	23.11.9	2) G	В	Published With international search report.		
(71) Applicant (for all designated States except US): LABORATORIES LIMITED [CA/CA]; 17 enue West, Willowdale, Ontario M2R 3T4	755 Ste					
<ul> <li>(75) Inventors/Applicants (for US only): CHONG, 32 Estoril Street, Richmond Hill, Ontario I THOMAS, Wayne [AU/AU]; 31 Taylor R W.A. 6009 (AU). YANG, Yan-Ping [CN/C 120 Torresdale Avenue, Willowdale, Onta (CA). LOOSMORE, Sheena [CA/CA]; 70 Drive, Aurora, Ontario LAG 4R4 (CA). SL Charles [CA/CA]; 189 Mabley Crescent, Th LAJ 2Z7 (CA). KLEIN, Michel [CA/CA]; 16 vard, Willowdale, Ontario M2P 1B9 (CA).</li> <li>(74) Agent: STEWART, Michael, I.; Sim &amp; M University Avenue, Suite 701, Toronto, On (CA).</li> <li>(54) Title: HAEMOPHILUS OUTER MEMBRANI</li> </ul>	LAC OF Road, N CA]; Ap ario M Crawfo A, Dwe cornhill, 6 Munr McBurn tario M	36 (CA) Redland St. 1709 Redland Rose Rose Rose Rose Rose Rose Rose Rose	). s, 9, 7 & a, 0 > 0			
(57) Abstract				DIS SEQUENCE CONFIREISON		
Purified and isolated nucleic acid from specific strains of Haemophilus influenzae is provided which encodes at least a portion of the D15 outer membrane protein of Haemophilus. The nucleic acid is used to produce peptides, polypeptides and proteins free of contaminant associated with Haemophilus for purposes of diagnosis and medical treatment. Furthermore, the nucleic acid may be used in the diagnosis of Haemophilus infection. Antisera obtained following immunization with the nucleic acid D15 outer membrane protein or peptides also may be used for the purpose of diagnosis and medical treatment.	SVIPIE	ALICOLONIC GAGTEGOST ANLÆRGAGS	erk.c	DALDERING AND	ATRICES VSSSTILED GELOPD  MELEPLES ALKENDITY RESULVA	Ca Eagun Huma Sauli Saul
	муж	Treswitte	FPW	NISINEGIGGERANGANELITGANSIATRANGANSALIYANGASITA	T THE THE THE THE THE THE THE THE THE TH	Ca Eagen Hirrin SBIJ PAK Ca Eagen Hirrin SBIJ SBIJ RMK

PT#VSDC9/IV/ITS\_FVONSMINESDRILESDAUGLEDAUGLEDAUGLEDAUGLEDAUGLEGO (SPLUFSA) (PROPERTIEST (PROPERT

Ca Eagun Hunna S233 FRK

30

35

## TITLE OF INVENTION HAEMOPHILUS OUTER MEMBRANE PROTEIN

#### FIELD OF INVENTION

5 The present invention is related to the field of molecular genetics and is particularly concerned with the cloning of an outer membrane protein D15 of Haemophilus.

#### BACKGROUND OF THE INVENTION

Haemophilus influenzae type b (Hib) is a major cause 10 of bacterial meningitis in children under the age of five years. Protective antibodies to the disease are induced by the capsular polysaccharide of the organism and a vaccine was developed that utilises the purified polyribosyl ribitol phosphate (PRP) as the antigen. This vaccine provides 90% protection in adults and in children 15 over 24 months of age, but was ineffective in children under 24 months Zangwill et al 1993 (The references are identified in a list of reference at the end of this disclosure). Like other polysaccharide antigens, PRP 20 does not induce the proliferation of T-helper cells, and re-immunisation fails to elicit either a booster response or an increase in memory cells. Conjugation of the PRP polysaccharide with protein carriers confers T-cell dependent characteristics to the vaccine substantially enhances the immunologic response to the PRP antigen. Currently, there are four PRP-carrier conjugate vaccines available. These are vaccines based upon <u>H. influenzae</u> type b capsular polysaccharide conjugated to diphtheria toxoid, tetanus toxoid, or Neisseria meningitidis outer membrane protein (reviewed in Zangwill et al 1993).

However, the current <u>Haemophilus</u> conjugate vaccines only protect against meningitis caused by Haemophilus influenzae type b. They do not protect against other invasive typeable strains (types a and c) and, more importantly, against non-typeable (NTHi) strains which are a common cause of postpartum and neonatal sepsis,

10

15

20

25

3

otitis media, epiglottitis, pneumonia, and tracheobronchitis, are required.

#### SUMMARY OF THE INVENTION

The present invention is directed towards the provision of purified and isolated nucleic acid molecules comprising at least a portion coding for a D15 outer membrane protein of a species of <a href="Haemophilus">Haemophilus</a>. The nucleic acid molecules comprising at least a portion coding for D15 outer membrane protein are useful for the specific detection of strains of <a href="Haemophilus">Haemophilus</a>, and for diagnosis of infection by <a href="Haemophilus">Haemophilus</a>. The purified and isolated nucleic acid molecules, such as DNA comprising at least a portion coding for D15 outer membrane protein, are also useful for expression of the D15 gene by recombinant DNA means for providing, in an economical manner, purified and isolated D15 outer membrane protein.

The D15 outer membrane protein or fragments thereof or analogs thereof are useful immunogenic compositions for the preparation of vaccines against diseases caused by <u>Haemophilus</u>, the diagnosis of infection by <u>Haemophilus</u> and as tools for the generation of immunological reagents. Mono- or polyclonal antisera (antibodies) raised against the D15 outer membrane protein produced in accordance with aspects of the present invention are useful for the diagnosis of infection by <u>Haemophilus</u>, specific detection of <u>Haemophilus</u> (in, for example, <u>in vitro</u> and <u>in vivo</u> assays) and for the treatment of diseases caused by infection by Haemophilus.

Peptides corresponding to portions of the D15 outer

membrane protein or analogs thereof are useful immunogenic compositions for the preparation of vaccines against disease caused by <a href="Haemophilus">Haemophilus</a>, the diagnosis of infection by <a href="Haemophilus">Haemophilus</a> and as tools for the generation of immunological reagents. Mono- or polyclonal antisera raised against these peptides, produced in accordance with aspects of the present invention, are useful for the

15

20

25

comprises at least an 18 bp fragment selected from the DNA molecules as recited above is inserted. The recombinant plasmid may be plasmid DS-712-2-1 having ATCC accession number 75604, deposited November 4, 1993 and plasmid JB-1042-5-1 having ATCC accession number 75006, deposited November 4, 1993.

The plasmids may be adapted for expression of the encoded D15 outer membrane protein in a host cell, which heterologous or homologous a incorporation into a recombinant vector, provided in accordance with a further aspect of the invention. recombinant vector may comprise at least a DNA segment comprising at least an 18 bp fragment selected from the DNA molecules as recited above and expression means operatively coupled to the DNA segment for expression of the gene product encoded thereby in the host cell. plasmid for expression of the encoded D15 outer membrane protein may be plasmid DS-880-1-2 having ATCC accession number 75605, deposited November 4, 1993 being adapted for expression at the D15 outer membrane protein in E. coli. The selected DNA segment may encode a polypeptide of at least 6 residues and, in particular, may be selected from those segments encoding a polypeptide of Table 2 (below). The DNA segment may further comprise a nucleic acid sequence encoding a leader sequence for export of the gene product from the host. The host for expression may be selected from, for example, Escherichia Bacillus, Haemophilus, fungi, yeast baculovirus expression system may be used.

Additional aspects of the invention include the protein encoded by the DNA molecule comprising at least a portion coding for the D15 outer membrane protein, fragment or a functional analog of such protein, the use of the protein or analog in vaccination and diagnosis, and the generation of immunological reagents. The invention also includes antisera (antibodies) raised

10

20

25

30

of the immunogenic composition or the nucleic acid molecule as recited above to provide protective immunity against <u>Haemophilus</u> infection.

The present invention further includes a chimeric molecule comprising a D15 protein corresponding thereto as provided herein linked to another polypeptide or protein or a polysaccharide. linked polypeptide or protein may comprise a surface protein or peptide corresponding thereto pathogenic bacteria, which may be the P1, P2 or P6 outer membrane protein of H. influenzae. The linked polysaccharide preferably comprise a PRP molecule from H. influenzae.

#### BRIEF DESCRIPTION OF THE FIGURES

The present invention will be further understood from the following description with reference to the drawings, in which:

Figure 1A shows the nucleotide sequence of the D15 gene from <u>H. influenzae</u> type b Ca strain (SEQ ID NO: 1) and its deduced amino acid sequence (SEQ ID NO: 2);

Figure 1B shows the nucleotide sequence of the D15 gene from <u>H. influenzae</u> type b Eagan strain (SEQ ID NO. 3) and its deduced amino acid sequence (SEQ ID NO: 4);

Figure 1C shows the nucleotide sequence of the D15 gene from <u>H. influenzae</u> type b MinnA strain (SEQ ID NO. 5) and its deduced amino acid sequence (SEQ ID NO: 6);

Figure 1D shows the nucleotide sequence of the D15 gene from <u>H. influenzae</u> non-typeable SB33 (SEQ ID NO. 7) and its deduced amino acid sequence (SEQ ID NO: 8);

Figure 1E shows the nucleotide sequence of the D15 gene from H. influenzae non-typeable PAK 12085 (SEQ ID NO. 9) and its deduced amino acid sequence (SEQ ID NO: 10);

Figure 1F shows an alignment of the nucleotide 35 sequences of the D15 genes (SEQ ID NOS: 1, 3, 5, 7 and 9)

10

15

20

30

35

97kDa); 2, GST standard; 3, GST-(D15 fragment) fusion protein; 4, fusion protein cleaved by thrombin; 5, N-terminal rD15 fragment; 6, GST; 7, low molecular weight markers;

Figure 10 shows guinea pig IgG antibody response to N-terminal rD15 fragment. The arrows indicate the immunization schedule. Bleeds were taken at 2, 4, 6 and 8 weeks. The bars represent the standard deviation; and

Figure 11 shows the hydrophilicity plot of D15 established by using a window average across 7 residues according to Hope, 1986.

#### GENERAL DESCRIPTION OF THE INVENTION

Any <u>Haemophilus</u> strains that have D15 genes may be conveniently used to provide the purified and isolated nucleic acid molecules (which may be in the form of DNA molecules), comprising at least a portion coding for a D15 outer membrane protein as typified by embodiments of the present invention. Such strains are generally available from clinical sources and from bacterial culture collections, such as the American Type Culture Collection. <u>H. influenzae</u> strains may include types a, b and c strains, non-typeable strains and other bacteria that produce a D15 protein, fragment or analog thereof. Appropriate strains of <u>Haemophilus</u> include:-

25 <u>H. influenzae</u> type b strain Ca;

H. influenzae type b strain MinnA;

H. influenzae type b strain Egan;

H. influenzae non-typeable b strain SB33; or

H. influenzae non-typeable b strain PAK 12085.

In this application, the term D15 outer membrane protein is used to define a family of D15 proteins which includes those having naturally occurring variations in their amino acid sequences as found in various strains of, for example, <u>Haemophilus</u>. The purified and isolated DNA molecules comprising at least a portion coding for D15 outer membrane protein of the present invention also

10

described herein are advantageous as diagnostic reagents, antigens for the production of <u>Haemophilus</u>-specific antisera, for vaccination against the diseases caused by species of <u>Haemophilus</u> and for detecting infection by <u>Haemophilus</u>.

Reference will now be made in detail to the presently preferred embodiments of the invention, which together with the following Examples, serve to explain the principle of the invention. For clarity of disclosure, and not by way of limitation, the detailed description of the invention is divided into the following sections:

# (i) The DNA sequences coding for the outer membrane protein D15 from <u>H. influenzae</u> type b Ca strain.

15 A clone producing the outer membrane protein designated D15 of <u>H. influenzae</u> type b (Hib) was isolated by screening a genomic library with H. influenzae type b OMP-specific polyclonal antibodies as previously described by Berns and Thomas 1965; Thomas and Rossi 20 The DNA fragment encoding the D15 protein was isolated, subcloned into pUC19 to produce pUC19/D15 (Figure 2) and used to transform E. coli HB101 as described in Example 1. Plasmid DNA was prepared from two individual colonies of E. coli HB101 containing the 25 pUC19/D15 plasmid. Sequencing was performed on an ABI DNA sequencer model 370A using dye-terminator chemistry and oligonucleotide primers which had been synthesized on an DNA synthesizer model 380B, and purified chromatography. Nucleotide sequence analysis of the D15 30 gene revealed that it contains a putative promoter and an open reading frame encoding 789 amino acids (Figure 1A).

The first 19 amino acid residues of the translated open reading frame form a typical leader sequence as found in other <u>H. influenzae</u> type b outer membrane proteins, such as P1 and P2. The N-terminal sequence of immuno-affinity purified native D15 antigen was

15

20

heterologous proteins in E. coli. The T7 expression system is described in U.S. Patent 4,952,496. were, therefore, constructed which utilize the T7 system to express a mature D15 protein that contains additional methionine residue at the amino terminus. The D15 signal sequence was removed during this construction process. A full length recombinant D15 (termed rD15) was expressed in inclusion bodies which allow the D15 protein to be readily purified. The D15 genes from H. influenzae type b strain Ca and H. influenzae non-typeable SB33 strain have been expressed at high levels in E. coli using the T7 system to permit production of large quantities of rD15 protein. The construction of clone DS-880-1-2 which expresses the SB33 D15 gene is described herein (see Figure 4 and Example 5). The rD15 protein was immunologically similar to its native counterpart isolated from <u>H. influenzae</u> typeable and non-typeable strains (see below). Thus, rD15 may be used as a crossreactive antigen in a diagnostic kit to detect many, if not all, strains of H. influenzae and other bacteria that produce a D15 outer membrane protein or analog thereof. Alternatively, rD15 can be used as an antigen to specifically detect the presence of H. influenzae in a sample.

25 A truncated D15 fragment was expressed in E. coli as a fusion protein with glutathione S-transferase (GST), as described in Example 6. The construction was designed to express the N-terminal fragment of the D15 protein. fusion protein was expressed at high levels from a pGEX-2T construction and the N-terminal fragment was cleaved 30 from the GST carrier protein by treatment with thrombin. This procedure generated a molecule termed the N-terminal rD15 fragment which encompasses amino acids 63-223 of the This N-terminal rD15 fragment was highly D15 protein. 35 immunogenic and elicited protective antibodies against challenge with live H. influenzae.

15

20

25

35

urea (see Example 8). After dialysis against PBS to remove urea, more than 80% of the D15 protein remained soluble. This soluble rD15 antigen was used for the immunogenicity studies described below. From shake-flask experiments, it was estimated that about 10 mg of soluble rD15 protein was obtained from 1 L of E. coli bacterial culture. It is clear that growing the recombinant E. coli strains under optimised fermentation conditions significantly increase the level of rD15 production.

10 (vi) Immunogenicity of the full-length recombinant D15 protien (rD15).

The immunogenicity of the full-length rD15 protein was studied in guinea pigs and mice. Using the immunization protocols described in Figure 7, a 15  $\mu$ g dose of rD15 induced high IgG titers in guinea pigs when administered in the presence of either Freund's adjuvant or AlPO<sub>4</sub>. In the mouse dose-response study, the protein appeared to be immunogenic at a dose as low as 5  $\mu$ g in either Freund's adjuvant (Figure 8A) or AlPO<sub>4</sub> (Figure 8B).

The protective ability of rD15 against <u>H. influenzae</u> type b infection was examined in the infant rat model of bacteremia essentially as described by Loeb (1987). Thus, infant rats passively immunized with guinea pig anti-rD15 antisera were significantly less bacteremic than controls injected with pre-bleed sera, which is consistent with the previous report by Thomas et al. (1990).

(vii) Purification and characterization of the N-terminal rD15 fragment.

The truncated rD15 fragment corresponding to the N-terminus of the D15 protein (residues 22 to 223) as described in Example 6, was expressed in <u>E. coli</u> as a soluble protein fused to GST. The fusion protein (46 kDa) was readily extracted using phosphate buffered saline (PBS). Purification of the GST-D15 fragment fusion

15

20

25

30

35

the immunization protocols described in Figure 10, a 10  $\mu$ g dose of N-terminal rD15 fragment induced a good booster response in guinea pigs with almost all the adjuvants tested. The highest anti-D15 IgG titer was observed in the group of guinea pigs immunized with N-terminal rD15 fragment in Freund's adjuvant. The second best adjuvant was Titermax (CytRx Inc.). The other two adjuvants, TPAD4 (tripalmityl-Cys-Ser-Glu<sub>4</sub>) and AlPO<sub>4</sub> were equally potent.

## 10 (ix) Protective ability of the N-terminal rD15 fragment against H. influenzae type b challenge.

An <u>in vivo</u> challenge model for a assessing the protective abilities of antigen against diseases caused by <u>Haemophilus</u> is the infant rat model of bacteremia as described by Loeb 1987. The protective ability of the N-terminal rD15 fragment against <u>H. influenzae</u> type b challenge was examined in this rat model. As illustrated in Table 1, infant rats passively immunized with rabbit anti-N-terminal rD15 fragment antisera showed significantly lower bacteremia compared to those injected with pre-bleed sera.

Since passively transferred antisera against the N-terminal rD15 fragment were found to be protective in the infant rat model of bacteremia, it was of interest to identify the protective epitope(s) of this N-terminal rD15 fragment. The first nine overlapping peptides of the D15 protein as listed in Table 2 were chemically synthesized based upon the amino acid sequence derived from the sequence of the D15 gene from H. influenzae type b Ca (Figure 1). These synthetic peptides were assessed for their reactivities with either rabbit or guinea pig antisera raised against purified N-terminal rD15 fragment by ELISAs. As shown in Table 3, both guinea pig and rabbit antisera reacted with a cluster of D15 peptides, including peptides D15-P4 to D15-P8 encompassing residues 93 to 209 of the D15 primary sequence.

10

15

20

25

30

35

was completely blocked by the addition of this mixture of five peptides (Table 5, group #2, 106%,  $p = 0.53 \times 10^{-8}$ ). These results strongly indicate that a cocktail of D15 synthetic peptides may be used as immunogens to induce protective antibodies against <u>H. influenzae</u>.

### (x) Epitope prediction and peptide synthesis.

To map the immunodominant T-cell or B-cell epitopes of D15, overlapping synthetic peptides covering the entire D15 protein sequence (Table 2 - SEQ ID NO: 14 to 49) were synthesized using the t-Boc solid-phase peptide synthesis as described in Example 15. The peptides were chosen based on their high index of hydrophillic  $\beta$ -turns estimated by secondary structure prediction analysis (Figure 11). Such peptides are likely to be surface-exposed and antigenic. Peptides more than 25 residues in length were selected to better mimic native epitopes.

## (xi) Identification and characterization of immunodominant epitopes of D15 using synthetic peptides.

linear B-cell map the epitopes of overlapping synthetic peptides representing the entire sequence of D15 were individually coated onto ELISA plates and probed with several anti-rD15 antisera as described in Example 19. The results are summarized in Table 6. Mouse antisera raised against rD15 reacted with all D15 peptides, but the major epitopes were located within peptides D15-P8 (residues 180-209 - SEQ ID NO: 21), D15-P10 (residues 219-249 - SEQ ID NO: 23), D15-P11 (residues 241-270 - SEQ ID NO: 24), and D15-P26 (residues 554-582 - SEQ ID NO: 39), respectively. Rabbit anti-D15 antisera recognized only peptides D15-P4 (residues 93-122 - SEQ ID NO: 17), D15-P14 (residues 304-333 - SEQ ID NO: 27) and D15-P36 (residues 769-798 - SEQ ID NO: 49). Guinea pig antisera raised against rD15 reacted with peptides D15-P2 (residues 45-72 - SEQ ID NO: 15), D15-P4 (residues 93-122 - SEQ ID NO: 17), D15-P6 (residues 135-164 - SEQ ID NO: 19), D15-P8 (residues 180-209 - SEQ ID

10

15

20

25

30

35

cells in the immune system is to provide helper activity for eliciting high levels of antigen-specific antibodies following immunization. Antigens containing Th1 epitope(s) stimulate antigen-specific T-cells to produce high levels of IL-2 and IFN- $\gamma$ , whereas Th2 epitope(s) induce high levels of IL-4 expression. Th0 epitope(s) stimulate the synthesis of IFN- $\gamma$  and IL-4.

Little is known about the cellular immune response to outer membrane proteins of H. influenzae and its role in the protection against H. influenzae infection and diseases. To this end, the inventors performed studies of the cellular response elicited in mice following rD15 immunization. D15-specific T-cell epitopes determined using D15 peptides and T-cell lines obtained from five BALB/c mice immunized with rD15 (see Example 23). The lymphocyte proliferative responses of the D15specific T-cell lines to overlapping D15 peptides were determined in conventional cytokine assays as described The results summarized in Table 7. in Example 24. revealed that stimulation only with certain synthetic peptides elicited proliferative responses and the release of specific cytokines. Synthetic peptides corresponding to residues 114-143 (D15-P5 - SEQ ID NO: 18), 282-312 (D15-P13 - SEQ ID NO: 26) and 577-602 (D15-P27 - SEQ ID NO: 40), and 219-249 (D15-P10 - SEQ ID NO: 23), 262-291 (D15-P12 - SEQ ID NO: 25), 390-416 (D15-P18 - SEO ID NO: 31), 410-435 (D15-P19 - SEQ ID NO: 32) 554-582 (D15-P26 -SEQ ID NO: 39), 596-625 (D15-P28 - SEQ ID NO: 41), 725-750 (D15-P34 - SEQ ID NO: 47) and 745-771 (D15-P35 - SEQ ID NO: 48) were shown to be highly stimulatory for rD15specific BALB/c Th0 cells and Th1 cells, respectively. Therefore, these immunodominant T-cell epitopes can be used as autologous carriers for PRP, and/or OMP B-cell epitopes to enhance their immunogenicity. The Th1 cell epitopes identified above may be useful in the H.

15

20

25

30

linked to polysaccharides including PRP as synthetic glycopeptide or lipoglycopeptide conjugates to produce alternate vaccines. These vaccines can be used to immunize against diseases caused by H. influenzae when administered to mammals, for example, by the intramuscular or parenteral route, or when delivered using microparticles, capsules, liposomes and targeting molecules, such as toxins or fragments thereof, and antibodies, to cells of the immune system or mucosal surfaces.

(xiv) Utility of D15 as carrier protein for the production of glycoconjugates.

To determine whether D15 may serve both as a protective antigen and a carrier, D15-PRP conjugation experiments were performed as described in Example 14. The D15-PRP conjugates were found to be highly immunogenic in rabbits and able to elicit both anti-D15 and anti-PRP IgG antibody responses as judged by D15-specific ELISA and PRP-BSA immunoassay (Table 9). These results clearly demonstrate the practical utility of D15 as a carrier protein for glycoconjugation technology.

In preferred embodiments of the present invention, the carrier function of D15 can be generally utilized to prepare chimeric molecules and conjugate vaccines against pathogenic bacteria, including encapsulated bacteria. Thus, the glycoconjugates of the present inventions may be applied to vaccinations to confer protection against infection with any bacteria having polysaccharide antigens, including, for example, Haemophilus influenzae, Streptococcus pneumoniae, Escherichia coli, Neisseria meningitidis, Salmonella typhi, Streptococcus mutans, Cryptococcus neoformans, Klebsiella, Staphylococcus aureus and Pseudomonas aeruginosa.

In another embodiment, the carrier function of D15
may be used, for example, to induce immunity toward
abnormal polysaccharides of tumor cells, or to produce

10

15

20

25

30

35

rD15 protein and its fragments were found to cross-react immunologically with the native D15 antigen isolated from both typeable and non-typeable H. influenzae isolates and thus represent cross-reactive immunogens for inclusion in a vaccine against diseases caused by H. influenzae. Furthermore, Haemophilus convalescent serum recognized D15 purified from H. influenzae as described herein, rD15 and N-terminal rD15 fragment.

another embodiment, the present invention provides a gene coding for the outer membrane protein D15 from H. influenzae having the specific nucleotide sequences described herein or ones substantially homologous thereto (i.e. those which hybridize under stringent conditions to such sequences), for genetically engineering hybrids or chimeric proteins containing a D15 fragment fused to another polypeptide or protein or a polysaccharide, such as H. influenzae outer membrane proteins, for example, P1, P2, or P6 or PRP. result, the hybrids, chimeric proteins or glycoconjugates may have higher protectivity against H. influenzae than D15, or P1, or P2, or P6, or PRP alone.

Thus, D15 outer membrane protein can function both as a protective antigen and as a carrier in a conjugate vaccine to provide autologous T-cell priming, wherein the of the conjugate hapten part is the capsular polysaccharide moiety (PRP) of H. influenzae. carbohydrate conjugate can elicit antibodies against both PRP and D15, and thus should enhance the level of H. influenzae-related diseases, protection against especially in infants.

In another embodiment, the present invention comprises an essentially pure form of at least one protein or peptide containing an amino acid sequence corresponding to at least one antigenic determinant of D15, which peptide is capable of eliciting polyclonal antibodies against H. influenzae in mammals. These D15-

10

20

25

<u>catarrhalis</u>, <u>Staphylococcus aureus</u>, or respiratory syncytial virus, in the presence or absence of adjuvant.

The D15 peptides (Table 2) or any portion, variant or mutant thereof, can easily be synthesized either manually or with a commercially available peptide synthesizer, such as the Applied Biosystems Model 430A synthesizer.

It is clearly apparent to one skilled in the art, that the various embodiments of the present invention have many applications in the fields of vaccination, diagnosis, and treatment of diseases caused <u>Haemophilus</u> infections, and the generation of immunological reagents. A further non-limiting discussion of such uses is further presented below.

### 15 1. Vaccine preparation and use

Immunogenic compositions, suitable for use as vaccines, may be prepared from immunogenic D15 outer membrane protein, fragments or analogs thereof and/or peptides corresponding to portions of D15 as disclosed herein. The vaccine elicits an immune response which produces antibodies, including anti-D15 outer membrane protein antibodies and antibodies against D15 that are opsonizing or bactericidal. Should the vaccinated subject be challenged by Haemophilus, the antibodies bind to the D15 outer membrane protein and thereby inactivate the bacterium. Opsonizing and bactericidal antibodies represent examples of antibodies useful in protection against disease.

Vaccines containing peptides are generally well
known in the art, as exemplified by U.S. Patents
4,601,903; 4,599,231; 4,599,230; and 4,596,792; all of
which references are incorporated herein by reference.
As to any further reference to patents and references in
this description, they are as well hereby incorporated by
reference without any further notice to that effect.
Vaccines may be prepared as injectables, as liquid

10

15

20

25

30

35

However, suitable dosage ranges are readily determinable by one skilled in the art and may be of the order of micrograms of the D15 outer membrane protein, analog, fragment and/or peptides. Suitable regimes for initial administration and booster doses are also variable, but may include an initial administration followed by subsequent administrations. The dosage of the vaccine may also depend on the route of administration and varies according to the size of the host.

The nucleic acid molecules encoding the D15 outer membrane protein of the present invention may also be used directly for immunization by administration of the DNA directly, for example, by injection for genetic immunization or by constructing a live vector, such as Salmonella, BCG, adenovirus, poxvirus or vaccinia. A discussion of some live vectors that have been used to carry heterologous antigens to the immune system are discussed in, for example, O'Hagan (1992). Processes for the direct injection of DNA into test subjects for genetic immunization are described in, for example, Ulman et al. (1993).

The use of peptides in vivo may first require their chemical modification since the peptides themselves may not have a sufficiently long serum and/or tissue halflife. Such chemically modified peptides are referred to herein as peptide analogs. The term peptide analog extends to any functional chemical equivalent of a peptide characterized by its increased stability and/or efficacy in vivo or in vitro in respect of the practice of the invention. The term peptide analog is also used herein to extend to any amino acid derivative of the described peptides as herein. Peptide analogs contemplated herein are produced by procedures that include, but are not limited to, modifications to side chains, incorporation of unnatural amino acids and/or their derivatives during peptide synthesis and the use of

10

30

35

nitration with tetranitromethane to form 3 nitrotyrosine derivative.

Modification of the imidazole ring of a histidine residue may be accomplished by alkylation with iodoacetic acid derivatives or N-carbethoxylation diethylpyrocarbonate.

Examples of incorporating unnatural amino acids and derivatives during peptide synthesis include, but are not limited to, use of norleucine, 4-amino butyric acid, 4amino-3-hydroxy-5-phenylpentanoic acid, 6-aminohexanoic acid, t-butylglycine, norvaline, phenylglycine, ornithine, sarcosine, 4-amino-3-hydroxy-6-methylheptanoic acid, 2-thienyl alanine and/or D-isomers of amino acids.

#### Immunoassays

The D15 outer membrane protein, analog, fragment 15 and/or peptides of the present invention are useful as antigens in immunoassays, including enzyme-linked immunosorbent assays (ELISA), RIAs and other non-enzyme linked antibody binding assays or procedures known to the 20 art for the detection of anti-bacterial, Haemophilus, D15 and/or peptide antibodies. In ELISA assays, the D15 outer membrane protein, fragment or analogs thereof and/or peptides corresponding to portions of D15 outer membrane protein are immobilized onto a selected surface, for example, a surface exhibiting a protein affinity, 25 such as the wells of a polystyrene microtiter plate. After washing to remove incompletely adsorbed D15 outer membrane protein, analog, fragment and/or peptides, a nonspecific protein, such as bovine serum albumin (BSA) or casein, that is known to be antigenically neutral with regard to the test sample may be bound to the selected This allows for blocking of nonspecific adsorption sites on the immobilizing surface and thus decreases the background caused by nonspecific bindings of antisera onto the surface. Normally, the peptides

10

15

20

25

30

35

<u>Haemophilus</u> and other bacteria that have genes encoding D15 outer membrane proteins.

The nucleotide sequences comprising the sequence encoding the D15 outer membrane protein of the present invention are useful for their ability to selectively form duplex molecules with complementary stretches of other D15 genes. Depending on the application, a variety of hybridization conditions may be employed to achieve varying degrees of selectivity of the probe toward the other D15 genes. For a high degree of selectivity, stringent conditions are used to form the duplexes, such as low salt and/or high temperature conditions, such as provided by 0.02 M to 0.15 M NaCl at temperatures of between about 50°C to 70°C. For some applications, less stringent hybridization conditions are required such as 0.15 M to 0.9 M salt, at temperatures ranging from between about 20°C to 55°C. Hybridization conditions can also be rendered more stringent by the addition of increasing amounts of formamide, to destabilize the hybrid duplex. Thus, particular hybridization conditions can be readily manipulated, and will generally be a method of choice depending on the desired results.

In a clinical diagnostic embodiment, the nucleic acid sequences of the D15 outer membrane protein genes of the present invention may be used in combination with an appropriate means, such as a label, for determining hybridization. A wide variety of appropriate indicator means are known in the art, including radioactive, enzymatic or other ligands, such as avidin/biotin, which are capable of providing a detectable signal. In some diagnostic embodiments, an enzyme tag, such as urease, alkaline phosphatase or peroxidase, instead of a radioactive tag may be used. In the case of enzyme tags, colorimetric indicator substrates are known which can be employed to provide means visible to the human eye or

10

15

20

25

30

35

The pBR322 plasmid, or other microbial plasmid or phage must also contain, or be modified to contain, promoters which can be used by the microbial organism for expression of its own proteins.

In addition, phage vectors containing replicon and control sequences that are compatible with the host microorganism can be used as a transforming vector in connection with these hosts. For example, the phage in lambda  $GEM^{TM}$ -11 may be utilized in making recombinant phage vectors which can be used to transform host cells, such as  $E.\ coli\ LE392$ .

Promoters commonly used in recombinant DNA construction include the  $\beta$ -lactamase (penicillinase) and lactose promoter systems and other microbial promoters, such as the T7 promoter system. Details concerning the nucleotide sequences of promoters are known, enabling a skilled worker to ligate them functionally with plasmid The particular promoter used generally is a matter of choice depending upon the desired results. Hosts that are appropriate for expression of transferrin receptor genes, fragment analogs or variants thereof include E. coli, Bacillus, Haemophilus, Bordetella, fungi, yeast, or the baculovirus and poxvirus expression systems may be used.

In accordance with an aspect of this invention, it is preferred to make the D15 outer membrane protein, fragment or analog thereof by recombinant methods, particularly since the naturally occurring D15 protein as purified from culture of a species of <u>Haemophilus</u> may include undesired contaminants, including trace amounts of toxic materials. This problem can be avoided by using recombinantly produced D15 outer membrane protein in heterologous systems which can be isolated from the host in a manner to minimize toxins in the purified material. Particularly desirable hosts for expression in this regard include Gram positive bacteria which do not have

10

20

25

30

35

the invention. Although specific terms have been employed herein, such terms are intended in a descriptive sense and not for purposes of limitations. Immunological and recombinant DNA methods may not be explicitly described in this disclosure but are well within the scope of those skilled in the art.

#### EXAMPLES

Methods of molecular genetics, protein biochemistry, and immunology used but not explicitly described in this disclosure and these EXAMPLES are amply reported in the scientific literature and are well within the ability of those skilled in the art.

#### Example 1

This Example illustrates the cloning and sequencing of the D15 genes.

Genomic DNA was purified from the <u>Haemophilus</u> influenzae type b strain Ca by lysis of the bacteria with pronase and sodium dodecylsulphate followed by phenol extraction and isopropanol precipitation, according to Berns and Thomas, 1965. The DNA was then partially digested with EcoRI and the DNA fraction containing 6-10 kb fragments was isolated following electrophoresis in low-melting point agarose. These fragments were ligated into a lambda gtll Ampl vector (Thomas and Rossi, 1986) and cloned as a lysogen into E. coli strain BTA282. Recombinant clones were selected for their ampicillin resistance conferred by the vector. To identify clones producing H. influenzae type b antigen, the clones were replica-plated on nitrocellulose filters and duplicate colonies induced for expression by temperature switch to 42°C for 2 hours. Colonies were lysed by wetting the filters with 1% sodium dodecylsulphate (SDS). The filters were then placed into a chloroform-saturated atmosphere for 15 min. The filters were then assayed by colony radioimmuno-assay using a hyperimmune rabbit anti-H. influenzae type b antiserum absorbed with E. coli lysate

15

20

25

30

plasmid transformed into <u>E. coli</u> HB101. Recombinant bacteria were found to produce the expected M<sub>r</sub> 80 kDa <u>H. influenzae</u> type b antigen when examined by Western blotting. The insert DNA was then characterised by restriction endonuclease mapping. A 2.8 kb <u>HindIII-EcoRI</u> fragment was subcloned into pUC19 to generate plasmid pUC19/D15, which was transformed into <u>E. coli</u> HB101. The recombinant bacteria expressed a M<sub>r</sub> 80 kD protein recognized by D15-specific antibodies on Western blot analysis of <u>E. coli</u> lysates.

Plasmid DNA was prepared from two individual colonies of recombinant E. coli HB101 containing the pUC19/D15 plasmid using standard techniques. Oligonucleotide sequencing primers of 17-25 bases in length were synthesized on the ABI model Synthesizer and purified by chromatography using OPC cartridges obtained from Applied Biosystems Inc., used in accordance with the manufactures recommendations. Samples were sequenced using the ABI model 370A DNA Sequencer and dye terminator chemistry according to manufacturers' protocols. This sequence indicated that the D15 gene contains an open reading frame encoding for 789 amino acids, including a putative signal sequence (Figure 1). The derived amino acid sequence was found to contain the sequence of an internal peptide obtained by thrombin digestion of native D15 that had been chemically determined. The amino composition of D15 derived from the D15 gene sequence was comparable (within experimental error) to that of the native protein as determined by amino acid analysis.

#### Example 2

This Example illustrates the preparation of chromosomal DNA from <u>Haemophilus influenzae</u> strains Eagan, MinnA, SB33, and PAK 12085.

35

H. influenzae Eagan and PAK 12085 chromosomal DNAs were digested with Sau3A I (0.5 unit/10  $\mu$ g DNA) at 37°C for 15 minutes and size-fractionated by agarose gel electrophoresis. Gel slices corresponding fragments of 15-23 kb were excised and DNA electroeluted overnight in dialysis tubing containing 3 mL of TAE (40mM Tris-acetate, 1mM EDTA, pH 8.0) at 14V. The DNA was precipitated twice and resuspended in water before overnight ligation with EMBL3 BamH I (Promega). The ligation mixture was packaged using the Lambda in vitro packaging kit (Amersham) according to the manufacturer's instructions and plated onto E. coli NM539 The library was titrated, then amplified and stored at 4°C under 0.3% chloroform.

MinnA chromosomal DNA (10  $\mu$ g) was digested with 15 Sau3A I (40 units) for 2, 4, and 6 minutes then sizefractionated on a 10-30% sucrose gradient in TNE (20mM Tris-HCl, 5mM NaCl, 1mM EDTA, pH 8.0). Fractions containing DNA fragments >5 kb were pooled 20 precipitated. In a second experiment, chromosomal DNA (2.6  $\mu$ g) was digested with Sau3A I (4 units) for 1, 2, and 3 minutes and size- fractionated by preparative agarose gel electrophoresis. Gel slices containing DNA fragments of 10-20 kb were excised and DNA extracted by 25 a standard freeze/thaw technique. The size-fractionated DNA from the two experiments was pooled for ligation with BamH I arms of EMBL3 (Promega). The ligation mix was packaged using the Gigapack II packaging kit (Amersham) and plated on E. coli LE392 cells. The library was titrated, then amplified and stored at 4°C under 0.3% 30 chloroform.

SB33 chromosomal DNA (20  $\mu$ g) was digested with Sau3A I (40 units) for 2, 4, or 6 minutes and size-fractionated on a 10-30% sucrose gradient in TNE (20mM Tris-HCl, 5mM NaCl, 1mM EDTA, pH 8.0). Fractions containing fragments >5 kb were pooled. In a second experiment, SB33

10

15

20

25

compared with the amino acid sequence of the D15 protein of <u>H. influenzae</u> type b Ca (Figure 3).

Example 5

EXAMPLE 5

This Example illustrates the expression of rD15 protein in E. coli.

A 2.8 kb fragment <u>HindIII-Eco</u>RI was subcloned into pUC19 and this pUC19/D15 plasmid was transformed into <u>E. coli</u> HB101. Upon induction, the positive clones expressed an 80 kDa protein which was recognized by D15-specific antisera on Western blot analysis. A <u>HindIII-Pst</u> I fragment was also subcloned into pUC19 and shown to express a 67 kDa protein. According to the restriction map, this 67 kDa protein corresponded to a C-terminal truncated D15 protein. On Western blot analysis, this truncated D15 was still recognized by the D15-specific antisera.

Plasmids to express the D15 gene of the non-typeable strain SB33 in <u>E. coli</u> were constructed. Plasmid JB-1042-5-1 containing the SB33 D15 gene and its flanking regions, was digested with <u>Eco</u>R I and <u>Hind</u> III and the 3kb D15 insert subcloned into pUC to give plasmid pRY-60-1 (Figure 4). Appropriate oligonucleotides were synthesized to restore the native D15 sequence between the ATG codon of the expression plasmid pT7-7 and the BsrF I site within the D15 gene. These oligonucleotides had the following sequence:

Nde

- 5'- TATGGCACCTTTTGTGGCAAAAGATATTCGTGTGGATGGTGTTCAAGGTG

  ACCGTGGAAAACACCGTTTTCTATAAGCACACCTACCACAAGTTCCACTGAATCT

  ACTTAGAATCAACAAACCGAGCAAGTTTACCTGTTCGTG SEQ ID NO: 50

  35 TGGTTGTTTAGGCTCGTTCAAATGGACAAGCACGGCC-5'- SEQ ID NO: 51

  BsrF I
- Plasmid pRY-60-1 was digested with <u>Eco</u>R I and <u>Bsr</u>F I and the DNA fragment containing most of the D15 gene was purified. pUC was digested with <u>Eco</u>R I and <u>Nde</u> I and the

10

15

25

30

35

produced by transformed  $\underline{E}$ ,  $\underline{coli}$  was isolated by affinity purification on glutathione agarose.

#### Example 7

This Example describes alternative expression systems for rD15.

The D15 gene or fragments thereof are also expressed in E. coli under the control of other regulated promoters. The D15 gene or fragments thereof are expressed in the absence of the leader peptide, or in other cloning systems where toxicity of D15 expression to the host is not problematic. The gene or fragments thereof are synthesized de novo or by employing the polymerase chain reaction using suitable primers. genes are cloned into suitable cloning vectors or bacteriophage vectors in E. coli or other suitable hosts directly when toxicity can be avoided. Expression systems are Gram-positive bacteria (such as Bacillus species), pox virus, adenovirus, baculovirus, yeast, fungi, BCG or mammalian expression systems.

#### 20 Example 8

This Example illustrates the protocol for extraction and purification of rD15 from <u>E. coli</u> expression system.

The cell pellet from a 250 mL culture, prepared as described in Example 5, was resuspended in 40 mL of 50 mM Tris, pH 8.0, and disrupted by sonication (3 x 10 min, 70% duty circle). The extract was centrifuged at 20,000 x g and the resulting pellet saved. The initial pellet was re-extracted with 40 mL of 50 mM Tris, 0.5% Triton X-100, 10 mM EDTA, pH 8.0. The suspension was then sonicated for 10 minutes at 70% duty circle. The extract was centrifuged at 300 x g for 5 minutes. The resulting supernatant was centrifuged again at 20,000 x g for 30 min and the resulting pellet was saved. The pellet was resuspended in 50 mM Tris, 0.5% Triton X-100, 10 mM EDTA, pH 8.0. The suspension was then mixed with PBS/ 8 M urea to a final urea concentration of 6 M. The solution was

10

15

20

25

30

35

This Example illustrates the procedure used for N-terminal rD15 fragment purification from GST using Glutathione-Sepharose 4B affinity chromatography.

A thrombin-digested GST-(D15 fragment) sample, prepared as described in Example 10, was loaded onto a Glutathione-Sepharose 4B column (2 mL) equilibrated with PBS containing 1% Triton X-100. The run-through of the column containing the N-terminal rD15 fragment was saved. After washing the column with 20 mL of PBS, the affinity column was regenerated by removing GST using 50 mM Tris-HCl buffer, pH 8.0, containing 5 mM glutathione. The purity of rD15 fragment was analysed by SDS-PAGE (Figure 9, lane 5). This N-terminal rD15 fragment contains amino acids 63-223 of the D15 protein as a result of cleavage at the spacious thrombin site shown in Figure 1A.

### Example 12

This Example illustrates the protocol used for the purification of D15-specific polyclonal antibodies by affinity chromatography using GST-(D15 fragment) fusion protein.

The recombinant GST-(D15 fragment) fusion protein, prepared as described in Example 9, was conjugated to cyanogen bromide-activated Sepharose. The affinity column was then used to purify antibodies from a rabbit hyperimmune anti-H. influenzae type b antiserum. The affinity purified-antibodies were shown by immunoblotting to react with a 80 kDa component present in the lysates of E. coli transformed with pUC9/D15 and in the lysates of several typeable and nontypeable H. influenzae isolates. These results confirmed that the DNA segment encoding the D15 fragment of the fusion protein was part of the open reading frame of the D15 gene.

Similarly, antisera raised against the recombinant fusion protein (Example 9) or the purified N-terminal rD15 fragment (Example 11) reacted with the D15 protein produced by H. influenzae strains (Example 13).

15

20

25

30

35

in Example 17. The mean molecular size of the PRP molecules used for conjugation was determined as being approximately 20,000 Daltons. The conjugation was carried out without a linker molecule but may also be carried out with a linker molecule. A PRP/D15 molar ratio of approximately 7 was used to provide an excess of PRP hapten.

The PRP/rD15 conjugate was tested according to the protocol of Example 18 for immunogenicity in rabbits and elicited both primary and secondary anti-PRP IgG and anti-D15 antibody responses (Table 9). Rabbit anti-rD15-PRP antisera also strongly reacted with both native D15 and rD15 as judged by immunoblot analysis. These data indicate that rD15 can be used as a carrier protein in a conjugate vaccine. In addition, a rD15-PRP conjugate vaccine should ensure a more consistent protection against H. influenzae type b disease, particularly in infants, as a result of the additional homotypic protection provided by antibodies directed against the D15 protein.

#### Example 15

This Example describes the preparation of D15 peptides.

D15 peptides (Table 2) were synthesized using an ABI 430A peptide synthesizer and optimized t-Boc chemistry as described by the manufacturer, then cleaved from the resin by hydrofluoric acid (HF). The peptides were purified by reversed-phase high performance liquid chromatography (RP-HPLC) on a Vydac C4 semi-preparative column (1 x 30 cm) using a 15 to 55% acetonitrile gradient in 0.1% trifluoryl acetic acid (TFA) developed over 40 minutes at a flow rate of 2 mL/min. All synthetic peptides (Table 2) used in biochemical and immunological studies were >95% pure as judged by analytical HPLC. Amino acid composition analyses of

10

15

25

30

35

equilibrated with 0.2 M sodium phosphate buffer, pH 7.2, and eluted with the same buffer. Fractions were monitored for absorbance at 230 nm. The first major protein peak was pooled and concentrated in a Centriprep 30 to 2.2 mL. The amount of protein was determined using the Bio Rad protein assay, and was found to be 300  $\mu$ g/mL. The presence of PRP in the protein conjugate fraction was confirmed by the Orcinol test.

#### Example 18

This Example describes the protocol used for the production of anti-PRP antisera in animals using rD15-PRP conjugates.

Rabbits were immunized intramuscularly with rD15-PRP conjugates (Example 14) (5 to 50  $\mu$ g PRP equivalent) mixed with 3 mg AlPO<sub>4</sub> per mL, followed by two booster doses (half amount of the same immunogen) at 2 week intervals. Antisera were collected every 2 weeks after the first injection, heat-inactivated at 56°C for 30 minutes and stored at -20°C.

#### 20 Example 19

This Example illustrates the reactivity between D15 peptides and anti-peptide and D15-specific antisera using D15-specific and peptide-specific ELISAs.

Microtiter wells (Nunc-Immunoplate, Nunc, Denmark) were coated with 200 ng of purified rD15 or 500 ng of individual peptides in 50  $\mu$ L of coating buffer (15 mM Na<sub>2</sub>CO<sub>3</sub>, 35 mM NaHCO<sub>3</sub>, pH 9.6) for 16 hours at room temperature. The plates were then blocked with 0.1% (w/v) BSA in phosphate buffer saline (PBS) for 30 minutes at room temperature. Serially diluted antisera were added to the wells and incubated for 1 hour at room temperature. After removal of the antisera, the plates were washed five times with PBS containing 0.1% (w/v) Tween-20 and 0.1% (w/v) BSA. F(ab')<sub>2</sub> fragments from goat anti-rabbit, guinea pig, mouse, or human IgG antibodies conjugated to horseradish peroxidase (Jackson ImmunoResearch Labs Inc.,

15

20

25

30

35

the substrate tetramethylbenzidine (TMB) in  $\rm H_2O_2$  (ADI, Toronto). The reaction was stopped with 1N  $\rm H_2SO_4$  and the optical density measured at 450 nm using a Titretek Multiskan II (Flow Labs., Virginia). A standard anti-PRP antiserum of known titer was included as positive control. Assays were performed in triplicate, and the reactive titer of each antiserum was defined as the reciprocal of the dilution consistently showing a 2-fold increase in O.D. value over that obtained with the pre-immune serum (Table 9).

#### Example 21

This Example describes the protocol used for the production of D15-specific antisera using purified D15, rD15 or N-terminal rD15 fragment.

New Zealand White rabbits (Maple Lane) and guinea pigs (Charles River) were immunized intramuscularly (IM) with a 10  $\mu$ g dose of either affinity-purified native D15 (Example 13), recombinant D15 (Example 8) or N-terminal rD15 fragment (Example 11) emulsified in Freund's complete adjuvant (Difco). Animals were boosted on day 28 with another 10  $\mu$ g dose of affinity-purified D15 or rD15 or rD15 fragment emulsified in Freund's incomplete adjuvant and bled on day 42 via the marginal ear vein. D15-specific polyclonal antibodies were purified from this material as described in Example 12.

#### Example 22

This Example illustrates the protective activity of D15-specific antisera against <u>H. influenzae</u> type b challenge using the infant rat model of bacteremia.

Five-day old infant rats were inoculated subcutaneously (SC) on the dorsum with 0.15 mL of two different rabbit anti-N-terminal rD15 fragments. Pre-immune sera were used as negative controls. One day after immunization, the infant rats were injected intraperitoneally (IP) with 200 colony-forming units (cfu) of Haemophilus influenzae type b Minn A strain (0.1)

10

15

20

25

30

35

supernatant added to further expand and maintain the viability of the peptide-specific T-cells. After a further 6 day-incubation, the cells were washed three times, each time with 200  $\mu L$  of culture medium.

Each set of cultures was then stimulated with the corresponding concentrations (1, 10 and 100 µg per mL) of the peptide in the presence of  $2 \times 10^5$  irradiated (1,500) rad) BALB/c spleen cells in a final volume of 200 µL of culture medium. Sixty µL of supernatant were then removed from each microculture. The supernatants from each triplicate cultures set were pooled. All supernatants were assayed for IL-2, Interleukin-4 and Interferon-gamma Detections of IL-2 and IL-4 were performed using murine IL-2 and IL-4 ELISA kits purchased from Endogen Inc. (MA, USA) respectively. Assay of IFN- $\gamma$  was performed using a mouse IFN-γ ELISA kit supplied by Genzyme Corporation (MA, USA). Test culture supernatants were assayed at 1 in 5 dilution according to the manufacturers' instructions. The results obtained are set forth in Table 7.

#### Example 25

This Example describes the general procedure used for the production of murine D15-specific monoclonal antibodies.

BALB/c mice were immunized intraperitoneally with 20 to 50  $\mu g$  of the N-terminal rD15 fragment (Example 11) emulsified in Freund's complete adjuvant. Two weeks later, the mice were given another injection of the same amount of immunogen in incomplete Freund's adjuvant (IFA). Three days before the fusion, the mice were boosted again with the same amount of immunogen in IFA. Hybridomas were produced by fusion of splenic lymphocytes from immunized mice with non-secreting Sp2/0 myeloma cells as previously described by Hamel et al. (1987). D15-specific hybridomas were cloned by sequential limiting dilutions and screened for anti-D15 monoclonal

#### TABLE 1

## PROTECTIVE EFFECT OF PASSIVELY TRANSFERRED ANTI-N-TERMINAL RD15 FRAGMENT ANTIBODIES IN THE INFANT RAT MODEL OF BACTEREMIA'

cfu/0.1 mL blood					
Rabbit antisera	Pre-immune	Post-immunization	p value		
Rb#434	510 (6/6)2	6 (1/6)	<0.001		
Rb#435	910 (4/4)	6 (1/4)	<0.001		

Five-day old infant rats were passively immunized with 0.15 mL of rabbit anti-N-terminal rD15 fragment s.c. One day later, the infant rats were challenged with 200 cfu of  $\underline{\text{H. influenzae}}$  type b strain MinnA (0.1 mL, IP). The blood samples were taken from each rat 24 hours after the challenge and analysed for bacteria counts.

The parentheses indicate the number of rats found to be bacteremic out of the total number of rats challenged.

D15-P29	619-646	LWVVSAKASAGYANGFGNKRLPFYQTYT	42
D15-P30	641-666	FYQTYTAGGIGSLRGFAYGSIGPNAI	43
D15-P31	662-688	GPNAIYAEYGNGSGTGTFKKISSDVIG	44
D15-P32	681-709	KISSDVIGGNAIATASAELIVPTPFVSDK	4.5
D15-P33	705-731	FVSDKSQNTVRTSLFVDAASVWNTKWK	. 46
D15-P34	725-750	VWNTKWKSDKNGLESDVLKRLPDYGK	47
D15-P35	745-771	LPDYGKSSRIRASTGVGFQWQSPIGPL	4.8
D15-P36	769-798	GPLVFSYAKPIKKYENDDVEOFOFSIGGSF	

TABLE 4

INHIBITION OF ANTI-N-TERMINAL rD15 FRAGMENT ANTIBODY-INDUCED PROTECTION BY D15 PEPTIDES IN THE INFANT RAT MODEL OF BACTEREMIA

Group #	Antibody	cfu / 10 μL blood	cfu in each group/cfu in group #4 (control) (%)
1	Anti-D15 Ab + PBS	60 ± 120 (3/7)	3
2	Anti-D15 Ab + peptides	300 ± 240 (6/7)	. 13
3	Anti-D15 Ab + rD15	1,520 ± 1,280 (7/7)	64
4	PBS + peptides	$2,360 \pm 1,200 (6/7)$	100

One half mL of rabbit anti-N-terminal rD15 fragment antiserum (Anti-rD15 fragment Ab) was mixed with either nine D15 peptides (100  $\mu g$  of peptides D15-P2 to D15-P10, See TABLE 2) or with 600  $\mu g$  of N-terminal rD15 fragment at room temperature for 1 hr. Antiserum and peptides mixed with PBS were used as controls. Seven-day old infant rats were injected s.c. with 0.2 mL of the various preparations. After 24 h, the infant rats were challenged I.P. with 200 cfu of H. influenzae type b strain MinnA. The blood samples were taken at 24 h after the challenge. The numbers in parentheses indicate the number of animals that were bacteremic out of the total number of animals challenged. The level of bacteremia is expressed as the mean of values obtained from seven infant rats tested individually  $\pm$  one standard deviation (SD).

TABLE 6

REACTIVITY OF RABBIT, GUINEA PIG AND MOUSE ANTI-rD15 ANTISERA WITH D15 PEPTIDES

Peptide	Rabbit <sup>2</sup>	Reactive Titer <sup>1</sup> Guinea Pig <sup>3</sup>	Mouse*
D15-P1	_		
D15-P2		•	+
D15-P3	-	+++	+,
	•	-	+
D15-P4	<b>+</b>	+	+
D15-P5	•	-	+
D15-P6	-	+	+
D15- <b>P</b> 7	•	•	+
D15-P8	-	++++	++++
D15-P9	~	-	+
D15-P10		•	+++
D15-P11	-	•	+++
D15-P12	-	•	+
015-P13	-	<b>-</b> ·	+
015-P14	+++	+	+
015-P15	-	-	+
015-P16	- ·	•	+
015-P17	-	· <b>-</b>	+
015-P18	· ·	• • • • • • • • • • • • • • • • • • •	+
015-P19	-	•	+
015-P20		•	+
)15-P21	•	•	+
15-P22	-		+
15-P23	-	-	+
15-P24	-	-	+
15-P25	•	-	+
15-P26	-	•	+++
15-P27	<u>.</u> .	, <b>+</b>	+

TABLE 7
T-CELL STIMULATORY ACTIVITY OF D15 PEPTIDES

D15-P1	- - - - 13 -
D15-P3 25	- - - 13 -
D15-P4       - <td>- - 13 -</td>	- - 13 -
D15-P5       742       38,000         D15-P6       -       -         D15-P7       -       -         D15-P8       -       -         D15-P9       -       -         D15-P10       108       1,900         D15-P11       -       -         D15-P12       1,052       6,100         D15-P13       105       6,200         D15-P14       -       -         D15-P15       -       -         D15-P16       48       -         D15-P17       -       -	- 13 - -
D15-P6	13 - -
D15-P7	-
D15-P8	-
D15-P9 108 1,900 D15-P11	
D15-P10       108       1,900         D15-P11       -       -         D15-P12       1,052       6,100         D15-P13       105       6,200         D15-P14       -       -         D15-P15       -       -         D15-P16       48       -         D15-P17       -       -	· •
D15-P11 6,100  D15-P13 105 6,200  D15-P14	-
D15-P12       1,052       6,100         D15-P13       105       6,200         D15-P14       -       -         D15-P15       -       -         D15-P16       48       -         D15-P17       -       -	-
D15-P13 105 6,200  D15-P14	-
D15-P14	٠_
D15-P15 D15-P16 48	56
D15-P16 48 - D15-P17	-
D15-P17	-
	-
D15-P18 32 4,800	-
	-
D15-P19 882 24,500	-
D15-P20	
D15-P21	-
D15-P22	•
D15-P23 78 -	•
D15-P24 103 -	-
D15-P25	<u>-</u>
D15-P26 572 6,700	•
D15-P27 274 7,505	68

67

TABLE 8

RABBIT AND GUINEA PIG ANTIBODY RESPONSES TO D15 PEPTIDES

	Peptide-s	Peptide-specific ELISAs		
_	Reactive	Titer		
Immunogen	Rabbit <sup>2</sup>	Guinea Pig³		
D15-P1	102,400	819,200		
D15-P2	204,800	1,637,400		
D15-P3	51,200	1,637,400		
D15-P4	204,800	819,200		
D15-P5	51,200	1,637,400		
D15-P6	51,200	409,600		
D15-P7	204,800	819,200		
D15-P8	51,200	409,600		
D15-P9	102,400	409,600		
D15-P10	102,400	819,200		
D15-P11	51,200	819,200		
D15-P12	102,400	204,800		
D15-P13	NT <sup>4</sup>	204,800		
D15-P14	NT	409,600		
D15-P15	NT	204,800		
D15-P16	NI	819,200		
D15-P17	NT	204,800		
D15-P18	NT	312,500		
D15-P19	nt	312,500		
D15-P20	NT	62,500		
D15-P21	NT	62,500		
D15-P22	NT	12,500		
D15-P23	NT	1,562,500		
D15-P24	NT	312,500		
D15-P25	NT	62,500		

TABLE 9

RABBIT IGG ANTIBODY RESPONSE TO D15-PRP CONJUGATE

		Reactive	Titer Against	2
Rabbit¹ #	PRP		rD15	
	2 doses	3 doses	2 doses	3 doses
489-1	1,600	3,200	1,600	6,400
490-1	1,600	1,600	6,400	25,600

Rabbits were immunized intramuscularly with rD15-PRP conjugates (5 to 50 μg PRP equivalent) mixed with 3 mg ALPO, per mL, followed by two booster doses (half amount of the same immunogen) at 2 week intervals.

Reactive titres is based on PRP specific and D-15 specific ELISAs.

71

#### **REFERENCES**

O'Hagan (1992).

Ulman et al., 1993.

Berns C.A. and Thomas C.A. (1965) J. Mol. Biol. 11:476-490.

Thomas W.R. and Rossi A.A. (1986) Infect. Immun. **52**:812-817.

Thomas W.R. et al. (1990) Infect. and Immun. 58:1090-1913.

Carlone G.M. et al. (1986) J. Clin. Microbiol. 24:330-331.

Smith, D.B. and Johnson K.S. (1988) Gene 67:31-40.

Harkness, R. et al. (1992) J. Bacteriol. 174:2425-2430.

Hamel et al. (1987) J. Med. Microbiol. 23:163-170.

Mills et al. (1993) Infect. Immun. 61:399-410.

Trinchieri, (1993) Immunol. Today 14:335-338.

Hope, T.P. (1986) J. Immunol. Methods 88:1-18.

Zangwill et al., 1993. MMWR 42:1-15.

Loeb et al. 1987. Infect. Immun. 55:2612-2618.

Panezutti, 1993. Infect. Immun. 61: 1867-1872.

- 8. The recombinant vector of claim 6 wherein said DNA segment encodes a polypeptide of at least 6 residues.
- 9. The recombinant vector of claim 8 wherein said polypeptide is selected from those shown in Table 2.
- 10. The recombinant vector of claim 6, 7, 8 or 9 wherein said DNA segment consists of no more than the coding sequence for said D15 outer membrane protein.
- 11. The recombinant vector of claim 10, wherein the DNA segment further comprises a nucleic acid sequence encoding a leader sequence for export of said gene product from said host.
- 12. A purified and isolated protein encoded by the DNA fragment contained in the recombinant vector of claim 10 or 11.
- 13. A purified and isolated D15 outer membrane protein, or a portion thereof.
- 14. The protein of claim 13 wherein the D15 outer membrane protein is a <u>Haemophilus</u> D15 outer membrane protein.
- 15. The protein of claim 14 wherein the D15 outer membrane protein is a <u>Haemophilus influenzae</u> D15 outer membrane protein.
- 16. The protein of claim 15 wherein the <u>Haemophilus</u> influenzae is a type b <u>Haemophilus influenzae</u> strain.
- 17. The protein of claim 16 wherein the <u>Haemophilus</u> influenzae type b strain is selected from Ca, MinnA and Eagan strains.
- 18. The protein of claim 15 wherein the <u>Haemophilus</u> influenzae is a non-typeable <u>Haemophilus</u> influenzae strain.
- 19. The protein of claim 18 wherein the non-typeable Haemophilus influenzae strain is selected from PAK12085 and SB33 strains.
- 20. A synthetic peptide containing an amino acid sequence corresponding to the amino acid sequence of the protein or portion thereof claimed in any one of claims

- 30. The chimeric molecule of claim 29 wherein said another polypeptide or protein comprises a P1, P2 or P6 outer membrane protein of <u>H. influenzae</u>.
- 31. The chimeric molecule of claim 28 wherein said polysaccharide comprises a PRP molecule from  $\underline{\text{H.}}$  influenzae.

FIG.1A

H. influenzae b Ca strain D15 sequence

			1/82			·
į	4 T 60	G A 120	ASP LGACT	r A TG 240	G C 300	, L T A 360
	T T	THR A.C.G	ASP G A (	ASN A A T (2	ALA G C G	ASP VAL L SATGTTA 360
0	AAT	THR ACA	G G T	ASP G A C	LYS ALA FAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	ĄSP GAT
10	r A T	THR A C G	GIN GLY CAAGGT	THRACTG	VAL G T G	SER TCAG
1	2 I C	GLY 3 G T A 110		WAL G T G 1 230	ASP G A T G 290	1LE A T T T 350
	ν Ν	PHE T T C (	GLY VAL GIN 3 G T G T T C A A 170	ARG	ASP 3 A T	ILE A T C A
 		LEU FTA7		GLY GLN 5 G T C A G (20)	PHE I T C G	LYS SER AAATCGA 140
! !	G <u>ттт G C A</u> ттаттаат G Атттта C G T C <u>татаат</u> ттат 30 40 50 60	RBS MET LYS LEU LEU LEU ILE ALA SER LEU LEU PHE GLY THR	ARG VAL ASP GTGTGGAT 160		SP VAL ALA ASN ILE VAL ARG   SER LEU PHE VAL SER GLY ARG PHE ASP VAL A TGTGCTA A TATTGTCGTTATTCGTA A GTGGTCGATTGATGTGTG 260 250 250	
, , ,	7	SER AGTI	( )	ALA G C C	GGT (	ALA G C T A
i	T T A L	ALA G C A A	ASP ILE SATATT(	ARG C G T G	SER AGTG	SER VAL VAL GCGTTGTG
1	T A T O	ILE 1 A T C	ASP G A T	WAL G T T	VAL G T A 10	VAL GTT
1	T T T T T T T T T T T T T T T T T T T	LEU CTA	LYS A A A G G 150	PRO V C C T G 210	PHE TTCG	SER A G C G 330
-35	TTG	LEU CTT	ALA G C A	LEU TTA(	LEU T T A	VAL G T T
		LYS A A A	AL T G	SER AGTT Mbinsi	SER T C T	VAL GTT(
i	20 g g	LYS A A A A 80	PHE T T T 140	ARG ALA SER LES CGAGCAAGTTT 200 spurious thrombin site	ILE VAL ARG   TTGTCCGCT 260	LEU C T T 320
	T A A	MET A T G	ALA ALA PRO   PHE 3 C C C C A C C T T T T 130 140	ARG C G A puriou	WAL G T C	ASP VAL LEUS ATGTGCT' 32
Hind III	- မ ပ	T C G	ALA G C A	ILE ATC s	ILE A T T	ASP G A T
<b>H</b>	G A T T A C G C C A A G C T T A A C G G T 10 20	RBS ATAGGATACAATC 70 start truncated GST/D15	HR VAL PHE ALA ALA PRO PHE V CTGTGTTTGCCGCACCTTTG 130	EU GLU GIN GIN ILE ARG ALA TAGAACAACAAATCCGAGCAA 190 spurious throm	ASN A A T 1 250	IS GIN GLU GLY ATCAAGAAGGCO
1	ပ ဗ	A 'F A	VAL PHE	GLN C A A	VAL ALA	GW GAA
	T A	RBS A G G	VAL 3 T G	GLU G A A	VAL G T G	GLN C A A
	<b>8</b>	A T I	氏 で T の	EU TA	SP A T G	IS A T

			2/82			
ဗ်ဗ္	D A C	JUQ		<sup>ဗ</sup> ဗ	G C G 720	н 4 0
ILE PRO THR GLU ALA LEU LYS GIN ASN LEU ASP ALA ASN ( TTCCCACTGAAGCACTTAAACAAACTTAGATGCTAACG 390 420	VAL L G T A A 480	ARG TYR ASN ALA THR VAL GLU PRO ILE VAL ASN THR L CGCTATAACGCAACAGTTGAACCTATTGTCAATACGC 510 520 530	T T	ASN GLU SER VAL SER SER THR LEU GLN GLU GLN MET G NACGAATCTGTTAGTAGCAGTACATTACAAGAACAAATGG 630 640 650	PE TTC	IU IYS ASP LEU GIN SER ILE ARG ASP TYR TYR LEU ASN ASN GLY TYR ALA LYS ALA GIN I AGAAAGATTGCAGTCAATTCGTGATTATTTAATTAATGGCTATGCCAAAGCACAAA 730 770 740 750 780
ALA G C T	LYS SER AAAGT	ASN A A T	ALA LYS 3 C A A A A	GLN C A A	GIN C A A	ALA G C A
ASP GAT(	LYS A A A	VAL GTC	ALA G C A	GLU GAA	ALA G C G	LYS A A A
LEU TTA 410	ALA G C C A 470	ILE ATTG 530	LYS A A A G 590	GLN C A A 650	GLY GGTG 710	ALA G C C 770
GIN ASN	GLU PHE 3 A A T T T	PRO C C T	ASP GATA	leu TTA	GU GAA	GLY TYR
GLN CAA	GLU GAA	GU GAA	ASP G A T	THR ACA	PHETTT	C G C
LYS A A A C 400	ASN A A T G 460	ARG TYR ASN ALA THR VAL GLU PRO GCTATAACGCAACAGTTGAACCTA 510 520	GLU GAA 580	SER A G T 640	LEU TRP GLY ASN LYS PHE GLU TTATGGGGAAATAAATTGAA(	ASN A A T G 760
THR GLU ALA LEU ACTGAAGCACTT 390	LEU TTA1	THR ACA	ASN A A T	SER AGCA 6	ASN A A T	ASN A A T I
ALA G C A	LYS A A A ?	ALA G C A	ILE A T C	SER AGT 2	G G A	LEU
t GLU TGAA 390	ARG GLU GAGAA 450	t ASN TAAC 510	ILE GIN ATTCAA 570	VAL GTTA	TRP T G G 90	TYR TYR 'ATTAT 750
THR ACT 33	ARG C G A	TYR TAT 5	ILE ATT S	SER TCTG 630	LEU TTA	TYR T
PROCCCA	ILE A T T	ARG C G C	LEU TTAA	GLU GAA	LYS A A A	ASP GATT
ILE A T T	LEU TTA	CLY GGT	ILE A T T	ASN A A C	TRP TGG1	ARG ♥ CGT(
SER VAL CTGTT 380	VAL G T T 1 440	VAL G T A 500	G A A 560	GC G G G 620	TRP T G G 680	ILE A T T 740
SER	ASP GAT(	SER A G T	ALA G C T	LYS A A G	PRO ASP SER C T G A T T C T 1 570	SER T C A
ASN A A C T	VAL GLX 3 T T G G C C	ALA G C A	ARG C G C	PHE TTC	ASP GAT	GIN C A G
GLY G G T 370	VAL G T T 430	HIS TYR ALA SER SACTATGCAAGT 490	ASN A A T 550	THR A C T 610	PRO C C T 670 GST/D]	LEU T T G 730
YS ILE LYS GLY ASN SER VAL A A A T C A A G G T A A C T C T G T T A 370 380	LY PHE LYS VAL GLY ASP VAL LEU ILE ARG GLU LYS LEU ASN GLU PHE ALA LYS SER GGTTTAAAGTTGGCGATGTTTAATTCGAGAAAATTAAATGAATTTGCCAAAAGT 430 440 450 450	YS GLU HIS TYR ALA SER VAL A A G A G C A C T A T G C A A G T G T A G 490 500	EU PRO ASN ASN ARG ALA TACCAAATAATCGCGCT 550	LA SER LEU THR PHE LYS GLY CATCATTAACTTTCAAGGGGA 610 620	AATTACAACCTGATTCTTGGT 670 end truncated GST/D15	LU LYS_ASP LEU AGAAGATTTG(
ILE A T C	PHE TTT1	YS GLU AAGAGO	PRO C C A	SER T C A	LEU TTA dtrun	LYS _
YS.	LY G G T	YS A A	EU T A	CA	LU A A en	IU A G

		3/8	32		
LEU ASN ASP CLU LYS THR LYS VAL ASN VAL THR ILE ASP V CTAAATGAAAAAAAAAGTTAATGTAACCATTGATG 810 820 830	ASP LEU ARG SER ALA ARG ILE ILE GLY ASN LEU GLY GLY M GACCTTCGTAGTGCACGCATTATAGGTAATCTGGGAGGTA 870 880 890	ARG S GTA 960	ASN ALA ILE LYS ALA LYS LEU GLY GLU ARG GLY TYR GLY S A A T G C A A T T A A A C T T G G A G A A C G C G G T T A C G G T A 990 1020	PRO ASP PHE ASP ALA ASN LYS THR LEU ALA ILE THR L CTGATTTTGATGAAAAAAACATTAGCGATAACC 1050 1060 1060 1070	ARG LEU THR VAL ARG GLN LEU ARG PHE GLU GLY ASN THR V CGTTTAACTGTTCGCCAACTTCGCTTTGAAGGAAATACCG 1110 1120 1140
ILE ATTG	CLX G G A (	LEU LEU SER ALA LEU HIS LEU ASN ASP THR PHE ARG TACTTTCAGCATTACATTAAATGATACTTTCCGCC 930 930	TYR TAC(	PRO ASP PHE ASP ALA ASN LYS THR LEU ALA ILE THR I CCTGATTTTGATGATGCAAATAAACATTAGCGATAACCC 1050 1050 1060	ASN A A T A
ASN ASP CLU LYS THR LYS VAL ASN VAL THR A A T G A T G A A A A G T T A A T G T A A C C A 810 820	LEU CTG	雅 T T C	G G T	ALA G C G	GLY G G A
VAL G T A 830	ASN A A T 890	THR A C T 950	ARG C G C 1010	LEU TTA 1070	GLU G A A 1130
ASN A A T	G G T	ASP G A T	LYS ALA LYS LEU GLY GLU PAAAGCAAAACTTGGAGAAC 1900	THR A C A	PHE TTT
VAL GTT	ILE	ASN A A T	GLY G G A	LYS A A A	ARG C G C
LYS A A A A 820	ILE ATT 880	LEU TTA 940	3 LEU A C T T 1000	A SN A A A T 1060	LEU CTT 1120
THR	ARG C G C	HIS C A T	LYS A A A	ALA G C A	GLN C A A
LYS	ALA G C A	LEU	ALA G C A	ASP G A T	ARG C G C
GLU F G A A 310	3 SER TAG1 870	ALA 1 G C A 130	I LYS T A A A 990	E ASP rrgar 1050	VAL 'GTT 10
ASP FGA7	ARG ICGI	SER LTCA	ILE A A T 1	PHE TTT	THR A C T
A A A 7	LEU	LEU	ASN ALA ILE ATGCAATT	ASP	LEU
				O	ARG C G T
THR ASP VAL GINACG ACCGATGTCAG	TYR 3 T A T 860	PRO A C C T 920	GLU A G A A 980	VAL 1 G T A 1040	ARG C G A 1100
VAL TGT	GLN A C A (	GLU FGA A	VAL rgta	SER LTCA	CLY GGA
ASP G G A '	LEU	LEU	ASP A. G. A. 1	ASN A A A 7	ALA
THR A A C (	GLY A G G 7 850	SER ALA GLU LEU GLU CTGCGAGCTTGA?	ALA F G C 1 970	ALA THR VAL ASN SER VAL CAACGGTAAATTCAGTA 1030	ASP r G A 1 1090
LYS T A A	GLU TGA	ALA I G C (	ILE	THR	WAL FGT7
LE THR LYS THR ASP VAL GIN TTACTAAAACGGATGTTCAG 790 800	AL ASN GLU GLY LEU GLN TYR TAAATGAAGGTTTACAGTAT 850 860		ER ASP ILE ALA ASP VAL GLU GTGATATTGCAGATGTAGAA 970 980	ER ALA THR VAL ASN SER VAL GCGCAACGGTAAATTCAGTA 1030 1040	EU VAL VAL ASP ALA GLY ARG TTGTTGATGCTGGACGA 1090
티	AL.	ET T G J	E 0	G E	T 1

			4	182			
	S T C	ی و	L A	S & S	g 8	T A O	S & 90
	37 S ATT 1200	ւ Մ 126	132 132	ГТА 1380	E 144	PHE T TTA 1500	'S ' A 156
	A A	S S	g S	II A T	II A T	至上	Z
	E E	# <b>4</b>	SV	<b>≒</b>	¥ <del>[</del>	er E	<b>4</b>
	T A	₽ A	EX A A	සු ව	S A	Σ, τ.	SE T C
	ط <sup>5</sup>	D 4	æ <b>€</b> -	<b>≈</b>	A	ွပ် ပ	<sub>ສ</sub> ັບ
	13 1 G	G A	TY A T	SE A	g ⊈	2 E	AS A A
	THR 1 A C T T 1190	္အ ပ	JO A 01	GLU S G A G A 1370	30 A	GLU 1 3 A G C 1490	P
	A C II	PH 17 1	48 T 2 13	GE 4	AC 24	GE 3 A 14	AS 3. A 15.
	× 4	မ္ကာပ	ن ت	× A	χ <sup>Ö</sup>	ت س	ک ک
	සි	器 🗗	A F	A C	AL G C	A C	A T
	5	> <del>[</del> -	<u>م</u> 4	> E	ق بر	۳ <del>۱</del>	ر ح
	ਰੇ <b>ਨ</b>	පි ස ස	AS G A	සි බ බ	9 B	TY A T	AS A A
	z	~ <del>«</del>	ی ق	ت ي	~ <del>«</del>	<u>ب</u> ج	A
	GLN A C A 7 1180	A C 240	VA 300	177 I. A 360	1H 4 C 120	G G G T 1480	GU 3 A 540
	_ A	, t	A	E	A A	, G C	. T
	E A	AR G	S A	GLY TYR GGTTAC 1360	ඩු	년 다	RH T
	ຼ ບ	. E-	, <b>€</b>	., E	_ o	<b>.</b> €	ູ
	A S	ASE 3 A	ASE 3 A	II T	ដ្	ASA A A	出土
	ຼິບ	A Q	<b>.</b> €		ູ້ບ	ຸ້ວ	, E-
	ME C	ET C	A A C	<del>g</del> 5	民ご	₽ ;	¥ € (
	SLU M A A A ' 1170	3 C 7 123(	T. 1 129(	РИЕ СП ТТГС С 1350	T 1	ER V. 3 T G ' 1470	J T ( 153(
	S A C	AR C	SE S &		ASI A A	SE C	AST A A
	ى ج	ω <del>Γ</del>	_ <del>-</del>	ے ت	- E	2 ک	_ ~ ~
	ARG GLN GLU MET ARG GLN GLN GLU GLY THR TRP TYR ASN S SGTCAGGAAATGCGCCAACAAGAAGGAACTTGGTATAATT 1170 1180 1190	LYS ILE ARG LEOU ASP ARG THR GLY PHE PHE GLU THR VAL GAAATCGTACAGGTTTCTTCGAAACAGTCG AAAATTCGCTTAGATCGTACAGGTTTCTTCGAAACAGTCG	ASN GLY SER ASN ASP GLU VAL ASP VAL VAL TYR LYS VAL L A A T G G T A G T G A T G T G G A T G T C G T A T A T A A G T C A T G T C G T A T A T A A G T C A T G T C G T A T A T A A G T C A T G T C G T A T A T A A G T C A T G T C G T A T A T A A A G T C A T G T C G T A T A T A A A G T C A T G T C G T A T A T A A A G T C A T G A T G T C A T G T C A T G T C A T G T C A T G T C A T G T C A T G T C A T G T C A T G T C A T G T C A T G T C A T G T C A T C T C A T	ILE ASN PHE CLY ILE CLY TYR CLY THR CLU SER CLY ILE S A T C A A C T T T G G T T A C G G T A C A G A G A G T G G T A T T A 1350 1360 1370 1380	ASI 3 A	是 C	ည် ည
	., E	, A	- E	ພິດ	, A		_ <b>E</b>
	\$ 50 C	CK A A	PA A	11. T 4	G A	၌ ဗ	<u>පි</u>
	£ 68	, <b>4</b> 7	္က ပင္ထဲ	₩ 19 7 19 7	.s & 00	% <del>1</del> %	ر <del>د</del> 8
	LEU T T A 1160	GLY 3 G A 1220	11. A T 12.	SER A G T 1340	LX 14 A	TYR F A T 1460	15.
	ALA ASP SER THR LEU SCTGATAGCACTTTA 1150 1160	5 4	C E	>₁ E-1	F	<u>ب</u> و	SER LEU GLY GLY ASN VAL PHE PHE GLU ASN TYR ASP ASN SER LYS S GTCTTGGTGGAAATGTTTTCTTTGAAACTACGATAACTCTAAAA 1520 1530 1560
	A TH	1 T	<b>E</b> C	aly G G T	₩ Đ	AS A B	SE SE
	<u>ت</u> ي	<sub>D</sub> o	٦ ←	ARG ASN THR CGTAACACGC 1330	~ ⊢	z	ر. آج
	es A G	ਲੇ ਨੂ	AS G A	A C	ES &	AS A A	₹ 5 T
	م ا	1	E E	z O	e e	S	. <del>.</del>
	ASF F G A 1150	VAL A G T 1210	IL A T 270	ASN F A A 1330	AL G C 390	1,YS 3, A, A 1,450	GC G G 510
	4 -	2 4 T	9 A 1	5 T	2 4	۳ <sub>و س</sub>	P T (
	G C	T T	C G	C G	ದ ೧	TH A C	AS G A
	SER CTC	z K	ر چ	D 4	κ <sub>E</sub>	<b>≻</b>	S
•	7 C	C A	AS A A	G A	YT A T	ය ල	LY:
	AL SER ALA ASP SER THR LEU P TTTCTGCTGATAGCACTTTAC 1150 1160	ER GLN LEU VAL GLU LEU GLY I CACAATTAGTTGAGTTAGGAA 1210 1220	LU ASN ARG ILE ASP PRO ILE ASN GLY SER ASN ASP GLU VAL ASP VAL VAL TYR LYS VAL 1 A A A A C C G A A T T G A T G G T A G T A A T G A T G A G T G G A T G T G	YS GLU ARG ASN THR GLY SER ILE ASN PHE CLY ILE GLY TYR CLY THR GLU SER CLY ILE S A A G A A C G T A A C G G G T A G T A T T G G T T A C G G T A C A G A G A G T G G T A T T A C G G T A C A G A G T G G T A T T A C B G T A G T G G T A T T A C B G T A G T G G T A T T A C B G T A G T G G T A T T A C B A C A C A C A C T T T G G T A T T G G T A C G G T A C A C A C A C A C A C A C A C A C A	ER TYR GLW ALA SER VAL LYS GLW ASP ASN PHE LEU GLY THR GLY ALA ALA VAL SER ILE A GTTATCAAGCAAGTGTTAAACAAGATAATTTCTTGGGAACAGGGGGGGCAGTAAGTA	LA GLY THR LYS ASN ASP TYR GLY THR SER VAL ASN LEU GLY TYR THR GLU PRO TYR PHE T CTGGTACGAAAATGATTATGGTACGAGTGTCAATTTGGGTTATACCGAGCCCTATTTA 1450 1450 1460 1470 1500	HP LYS ASP GLY VAL SER LEU GLY GLY ASN VAL PHE PHE GLU ASN TYR ASP ASN SER LYS ( C'TAAAGATGGTGTAAGTCTTGGAAATGTTTTCTTTGAAACTACGATAACTCTAAAA 1510 1520 1530 1530
	∢ ₽	ыO	<b>1</b>	׫	ភា ល	J	Ħ U

		5	/82	•		
ARG THR THR TYR GLY SER ASN VAL THR LEU GLY PHE P CGTACGACTTACGGAAGTAATGTTACTTTAGGTTTCC 1590 1620	TYR TYR VAL GLY LEU GLY HIS THR TYR ASN LYS ILE SER A TATTATGTAGGATTAGGTCATACCTATAATAAATTAGTA 1650 1680	ARG ASN LEU TYR ILE GIN SER MET LYS PHE LYS GLY ASN G CGTAATTTATATATTCAATCAATGAAATTAAAGGTAATG 1710 1720 1730	ASP PHE SER PHE GLY TRP ASN TYR ASN SER LEU ASN ARG GGATTGGAACTATAACAGCCTTAATAGAGATT80 1770 1780 1780	VAL LYS ALA SER LEU GLY GLY ARG VAL THR ILE PRO GLY S GTTAAAGCAAGTCTTGGTGGACGAGTTACTATTCCAGGTT 1830 1860	LEUSER ALA ASPVALGINGLYPHETYRPROLEUASPARGA CTAAGTGCAGATGTACAGGGTTTCTACCCATTAGACAGAG 1890 1920	ALA LYS ALA SER ALA GLY TYR ALA ASN GLY PHE GLY ASN L GCAAAAGCATCTGCAGGATATGCAAATGGTTTTGGAAACA 1950 1980
THR LEU GLY 1 C T T T A G G T 1	ILE A T T	G G T	ASN A A T	SER LEU CLY CLY ARG VAL THR ILE PRO CLY AGTCTTGGTGGACGAGTTACTATTCCAGGT' 1860 1860	ASP G A C	GLY G G A I
LEU TTA	ASN LYS ILE NATAAAATT 1670	LYS A A A	LEU	ILE A T T	LEU T T A	PHE TTT
THR A C T 1610	ASN A A T 1670	PHE T T T 1730	SER A G C 1790	THR ACT 1850	PRO C C A ' 1910	GLY G G T 1
VAL G T T	TYR T A T	LYS A A A	ASN A A C	VAL G T T	TYR T A C	ASN A A T
SER ASN VAL GTAATGTT1	VAL GLY LEU GLY HIS THR TYR STAGGATTAGGTCATACCTATA 1650	LEU TYR ILE GIN SER MET LYS FTATATATTCAATCAATGAAAT 1710	SER PHE GLY TRP ASN TYR ASN CTTTTGGTTGGAACTATAACA 1770	ARG C G A	ALA ASP VAL GIN GLY PHE TYR CAGATGTACAGGGTTTCTACC 1890	ALA G C A
SER A A G T 1600	HIS C A T 1660	SER T C A 1720	ASN A A C 1780	GLY G G A 840	GLY 3 G G T 1900	TYR T A T 960
GLY G G A	GLY GGT	GEN CAA	TRP TGG	GGT 1	GIN CAG	GLY GGA 1
THR TYR GLY ACTTACGGAA	LEU T T A	ILE A T T	G G T	LEU CTT	VAL G T A	ALA G C A
т тнк 3 G A C т 1590	GLY G G A 50	TYR T A T 10	PHE TTT 70	SER AGT 30	ASP GAT 90	SER TCT 50
THR A C G A 1590	VAL GTA 16	LEU TTA	SER TCT 17	ALA G C A A 1830	ALA G C A 189	ALA G C A 195
ARG C G T A	TYR TYR 'ATTAT	ARG ASN GTAATT	PHETT	LYS A A A G	SER AGTG	LYS A A A
LYS A A G (				VAL GTTP		ALA G C A
SER SER ASN TYR CCTCTAACTAT 70 1580	SER T C C 1640	ASN A A C 1700	PHE TTT 1760	GLY G G G 1820	LYS A A A 1880	SER T C T 1940
ASN F A A C	ASN A C	TYR 1 T A T	ASP	LYS	TYR TAC	VAL G T A
SER	ASN A A 7	GLU GAP	ASN A A T	TR	TYR TAC	VAL G T T
SER A T C ( 1570	GLU r G A A 1630	ALA LEU GLU TYR 3 C T C T A G A A T A T A	THR A A C A 1750	PRO C C A 1810	LYS C A A A 1870	LEU TRP VAL VAL SER TCTGGGTTGTATCT 1930 1940
ASP THR	ASN A A A 1	ALA G C 1	LYS	PE	ASN A A C	LEUCTC
ER ASP THR SER SER ASN TYR GTGATACATCCTCTAACTAT 1570 1580	RO VAL ASN GLU ASN ASN SER CTGTAAATGAAAATAACTCC 1630 1640	SN PHE ALA LEU GLU TYR ASN ACTTTGCTCTAGAATATAAC 1690 1700	LY ILE LYS THR ASN ASP PHE GCATTAAAACAAATGACTTT 1750 1760	LY TYR PHE PRO THR LYS GLY GCTATTCCCAACTAAAGGG	ER ASP ASN LYS TYR TYR LYS CTGATAACAAATACTACAAA 1870 1880	SP HIS LEU TRP VAL VAL SER ATCACCTCTGGGTTGTATCT 1930 1940
E O T O	& ၁	N S	G C	C C	က 1	SP A T

		6/8			
THR TYR THR ALA CLY CLY ILE CLY SER LEU ARG CLY PHE A CTTATACAGGGGGCATCGGTTCATTACGTGGTTTG 2010 2020 2030	ASN ALA ILE TYR ALA GLU TYR GLY ASN GLY SER GLY THR G A C G C A A T T T A T G C C G A A T A T G G T A G T G G T A C T G 2070 2070 2080	LY THR PHE LYS LYS ILE SER SER ASP VAL ILE GLY GLY ASN ALA ILE ALA THR ALA SER A GTACTTTTAAGAAGATAAGTTCTGATGTGATTGGTGGTAATGCAATCGCTACAGCTAGCG 2110 2120 2130 2130	LA GLU LEU ILE VAL PRO THR PRO PHE VAL SER ASP LYS SER GIN ASN THR VAL ARG THR S CAGAGTTAATTGTGCCAACTCCATTTGTGAGCGATAAGAGCCAAAATACGGTCCGAACCT 2170 2180 2190 2200	ER LEU PHE VAL ASP ALA ALA SER VAL TRP ASN THR LYS TRP LYS SER ASP LYS ASN GLY L CCTTATTTGTTGATGCGGCAAGTGTTTGGAATACTAAATGGAAATCAGATAAAAATGGAT 2230 2240 2250 2280	EU GLU SER ASP VAL LEU LYS ARG LEU PRO ASP TYR GLY LYS SER SER ARG ILE ARG ALA S TAGAGAGCGATGTATTAAAAGATTGCCTGATTATGGCAAATCAAGCCGTATTCGCGCCT 2390 2330 2340
ARG GLY	GGT 2	THR ALA	ARG C G A	ASN A A T (	ARG C G C G
ARG C G T	SER AGTO	THR ACA	THR VAL ACGGTC( 2210	ASP LYS 3 A T A A A 1 2270	ARG ILE GTATT( 2330
LEU T T A C 2030	GLY G G T 2090	ALA G C T A 2150	THR A C G ( 2210	ASP 3 A T 7 2270	ARG C G T 7 2330
SER	ASN A A T	ILE A T C (	ASN A A T 1	SER FCA(	SER AGC
GLY ILE GLY SER LEU GCATCGGTTCATTAC 2020 2030	ILE TYR ALA GLU TYR GLY TTTATGCCGAATATGGT? 2070	ALA G C A A	SER GIN AGCCAAP 200	TRP LYS SER CGGAAATCA(	ASP TYR GLY LYS SER SER SATTATGGCAAATCAAGC 2320
ILE NATC 2020	TYR \ T A T 2080	ASN 2140	SER 3 A G C 2200	TRP \ T G G . 2260	LYS 2320
GEY GGC 2	GLU GAA 2	VAL IIE GLY GLY 1 T G A T T G G T G G T 1 2130	LYS A A G A 22	LYS A A A T 22	GLY GGC, 2
THR ALA GLY 1 C A G C G G G T C 2010	ALA G C C	G G T	ASP GAT1	ASN THR	TYR T A T
ALA G C G 10	TYR TAT 70	ILE ATT 30	SER AGC	ASN A A T	ASP GAT .0
THR A C A G 2010	ILE T ATTT 2070	VAL G T G A 2130	VAL : G T G A 2190	TRP 1 T G G A 2250	PRO 7 C C T G 2310
TYR T A T	ALA G C A	ASP GAT	PHE TTT	SER VAL	LEU TTG
THRACTI		SER TCT (	PRO C C A 1	SER AGT	ARG A G A
GLN C A A 2000	PRO C C T 2060	SER A G T 2120	THR A C T ( 2180	ALA G C A 2240	LYS A A A 2300
PHE TYR GIN	G G A	ILE A T A	PRO C C A	ALA G C G	LEU TTA
PHE TTC	ILE	LYS A A G	VAL G T G	ASP G A T	VAL G T A
ARG LEU PRO GTTTACCGT 1990	SER ' A G T 2050	THR PHE LYS LYS ILE SER CTTTTAAGAAGATAAGT 2110	ILE ATT 2170	VAL G T T 2230	ASP C G A T 2290
LEU	GLY G G T	PHE	LEU	PHETTT	SER A G C
	LA TYR GLY SER ILE GLY PRO CTTATGGTAGTATTGGACCTA 2050 2060	THR'ACT	LA GLU LEU ILE VAL PRO THR CAGAGTTAATTGTGCCAACT 2170	ER LEU PHE VAL ASP ALA CCTTATTGTTGATGCGG 2230	EU GLU SER ASP VAL LEU TAGAGGGATGTATTA1 2290
YS A G C	CT	GT A	CA	ر د د	EU T A

TRP GLN SER PRO ILE GLY PRO LEU VAL PHE SER TYR ALA L TGGCAATCTCCTATTGGCCATTGGTATTCTCTTATGCA 2370 2380 2390 2400	ASN ASP ASP VAL GLU GLN PHE GLN PHE SER ILE GLY GLY S A A T G A T G T C C G A T T T A G T A T T G G A G G T T C C A A T T T A G T T G G A G G T T C C A A T T T A G T A G G T T C A A T T A G T A G A G G T T C C A A T T T A G T A G G A G G T T C C A A T T T A G T A G G A G G T T C C A A T T T G G A G G T T C C A A T T T G G A G G T T C C A A T T T G G A G G T T C C A A T T T G G A G G T T C C A A T T T G G A G G T T C C A A T T T G G A G G T T C C A A T T T G G A G G T T C C A A T T T G G A G G T T C C A A T T T G G A G G T T C C A A T T T G G A G G T T C C A A T T T G G A G G T T C C A A T T T G G A G G T T C C A A T T T A G T A T T G G A G G T T C C A A T T T A G T A T T G G A G G T T C C A A T T T A G T A T T G G A G G T T C C A A T T T A G T A T T G G A G G T T C C A A T T T A G T A T T G G A G G T T C C A A T T T A G T A T T G G A G G T T C C A A T T T A G T A T T G G A G G T T C C A A T T T A G T A T T G G A G G T T C C A A T T T A G T A T T G G A G G T T C C A A T T T A G T A T T G A T G T C C A A T T T A G T A T T G A T T T A G T A T T G T C A T T T A G T A T T T A G T A T T T A G T A T T T A G T A T T T A G T A T T T A G T A T T T A G T A T T T A G T A T T T A G T A T T T A G T A T T T A G T A T T T A G T A T T T A G T A T T T A G T A T T T A T T T A G T A T T T A T T T A T T T T
TYR /	GLY C
R T T T	ອີ ອີ
SER TC1	ILE ATT
PHE T T C 2390	SER A G T 2450
VAL G T A	PHETT
LEU TTG	GLN C A A
PRO C C A 2380	PHE T T C 2440
G G G	GLN C A G
ILE O	GLU GAA
PRO C C T	WAL GTC 0
SER TCT 237	ASP G A T 243
GLN C A A	ASP G A T
TRP T G G	ASN A A T
GLN C A A 2360	GLU 3 A A 2 2420
YE TTC	TYR FAT(
G G A	LYS A A A 7
VAL ' G T C ( 2350	LYS ' A A A 2
GLY 3 G T ( 2)	ILE ATT? 24
ER THR GLY VAL GLY PHE GLN CTACAGGTGTCGGATTCCAAT 2350 2360	YS PRO ILE LYS LYS TYR GLU AACCAATTAAAAAATATGAA 2410 2420
C T 7	YS A A C

	ТТТТСТТСАGAACTCAAAAACAACGTTCTGCCTAA 2490 2520 2510 2520
	ACAAC
	T C A A A A A 2500
	CAGAAC
	CTTCAT 2490
٠,	
	ААТТGААСТ1 2480
* *	CTTTCTAATAAA 2470
*	T A
呂	T T C
題	CT

82
A A T G A A 2580
G G A T A T T T A T C A 2570
T A A T T A A 2560
A T T A A A C C C A T C A T T 2550
G A A A A T A T T P 2540
T T A A T T G G G C A G A 2530
F

T G C T T C 2640
C T T C A G G C T A 2630
TATTGCACTTG 2620
T G C T T T A G G 2610
TAACCGCACT 2600
A C A T C G C A A A A G 2590
AA

	GATCGC	2700
	ACATCACCCA	2690
* C H H H H H W H W		2680
TGCAGGT		0/97
TGCTTTCATTAA	2660	0007
CGCTGAAGAAAAAT	2650	

8/82

ATAAAATTGCTGCTCGTAAAAAGTAGAAGCAAAGTT 80 2790 2820	; C A C C T C G C T C A A G C T G A T A T T C A A A A C G C C A A C A G 140 2850 2850 2860 2860	
ГАААТТ G С T G C T 2790	A C C T C G C T T A C G T 2850	•
;	G C T T T A G A A A A A G A T G 2830 28	
A G	ອ ວ ອ	

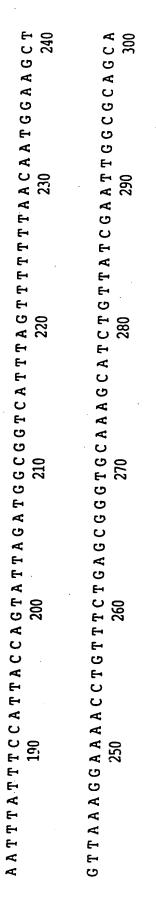
FIG.1A.(CONTINUED)

GATAAAAAA

FIG.1B

DS-712-2-1 DNA, Eagan D15 sequence IS THE SEQUENCE BEING TRANSLATED

9/82



CTGTTATTAAGCTTAACGGTGTTTGCATTATTTAATGATTTTACGTCTATAATATA 320

TAGGATA CAATCGATGAAAAA CTT CTAATCGCAAGTTTATTATTCGGTACGACACGAC	710 720
THR THR  S A C A A C G  GLY ASP  G G T G A C  G G T G A C  G A C A A T  G A C A A T  G A C A T  G A C G  ASP  ASP  ASP  ASP  G A C A A T  G A T G T T  G A T G T T	710
THR 3 A C A GLY 4 G G T G A C 1 A A A A 1 G A T 1 G A T 1 G A T 1 G A T 1 G A T 1 G C T 1	710
	710
THR GIN GIN CAAG ACTG ACTG CAAG ACTG ACTG ACTG ACTG	710
GG T A 410  VAL CG T C 470  VAL TG A 530  ILE S A T T T 650  LEU A T T A G T T A G	
LEU PHE GLY 41  ASP GLY VAL 3 A T G G T G T T C GLN ARG VAL C A T G G T G T G SER ILE ILE C G A T C A T T T G A A A A C T T A G A A A A C T T A G  T T A A A A C T T A G A A A A C T T A G A A A A A C T T A G  T T A A A A A C T T A G  T T A A A A C T T A G  T T A A A A C T T A G  T T A A A A C T T A G  T T A A A A C T T A G  T T A A A A C T T A G  T T A A A A C T T A G  T T A A A A C T T A G  T T A A A A C T T A G  T T A A A A C T T A G  T T A A A A C T T A G  T T T A T T C G A T C T T A G  T T T T T T T T T T T T T T T T T T	
TAGGATACAATCGATGAAAAACTTCTAATCGCAAGTTTATTATTCTTCTAATCGCAAGTTTATTATTTCTTCTAATCGCAAGTTTATTATTTCTTCTAATCGCAAGTTTATTATTTTCTTCTAATCGCAAGTTTATTATTTTTTTT	
SER LEU  G T T T A  400  ARG VAL  G T G T G  G T C G G T  C C G G T  S20  GT C G A  T T A A A  T T A A A C	700
ALA SER  G C A A G T  ILE ARG  ILE ARG  ILE ARG  ILE ARG  ILE ARG  ILE G C C  G T G C C  G T G C C  ILE ALA  ILE C A C T T  ILE ALA  ILE ALA  ILE ALA  ILE ALA  ILE ALA  ILE C A C T T  ILE ALA	
ALA GCAA ARG ARG CGTG AGTG AGTG ALA ALA GCAC	
ILE ATC (390 ASP GATP 450 VAL GTT G GT A A 570 CGT G GGA A G GGA A G	969
LEU ILE CTAAT 390 LYS ASP AAAGA' PRO VAL CCTGT 510 PHE VAL TTCGT7 570 SER VAL AGCGT7 AGCGT7	
INS LIS LEU LEU  380  PHE VAL ALA LYS  "TTGTGGCAAA  440  ALA SER LEU PRO CAAGTTTATT  500  LEU VAL VAL SER  TTGTTGTGTTAG  620  VAL ILE PRO THR  TTATTCCCACT	
LYS LEU A A A C T T 30 VAL ALA G T G C A 0 SER LEU T C T T T A 0 VAL VAL 0 LILE PRO ILLE PRO A T T C C C	<b>&gt;</b>
TAGGATACAATCGATGAAAAACTTCTA 370  380  VAL PHE AIA AIA PRO PHE VAL AIA LYS  TGTGTTGCCGCACTTTTGTGCCCAAAA  430  AGAACAAATCCGAGCAAGTTTACCT  SOO  VAL AIA ASN ILE VAL ARG SER LEU PRO  TGTGGCTAATTTGTCCGCCTCTTTACCT  SOO  TCAAGGCGATGTGCTTTATTATTC  TCAAGAACTTTACTCCGTTTTATTCTC  SOO  TIE LYS GTY ASN SER VAL ILE PRO THR  AATCAAAGGTAACTCTGTTATTCCCACT	080
MET ALA PRO ILE ARG TTCCGA TTTGTC TTTGTC ASP VAL ASP V	•
TCG A ALA GCAC ILE ATCC ATTG ASP SATG ASN	
AGGATACAAT 370 370 VAL PHE ALA GTGTTTGCCG 430 GLU GLN GLN GAACAAAA GAACAAAA GAACAAAA GTGCTAATA GTGCTAATA 550 GTGGCTAATA 550 GTGGCTAATA 550 GTGGCTAATA 550	2
VAL PHE TGTTTC GLU GIN AACAAC TGGCTP	
VAL GIU GIU GIU T G G T G G T G G T C G T C A	
T T B T B T B T B T B T B T B T B T B T	

GLN LEU ASN

ME.

ASP

THR LYS

		11	182		
LYS A A 780	LEU C T 840		_	GLU 3 G A 1020	ILE A T 1080
VAL GTAP	THR A C G	LEU TTG	MET A T G	PHE T'T C	GIN ILE CAAAT 1080
SER A G T	ASN A A T	LYS A A A '	GIN C A A	GIN C A A	ALA G C A
ALA LYS CCAAA 770	VAL 'GTC1	ALA G C A A 0	GIN GLU SAAGAAC 950	ALA G C G	ALA LYS CCAAA(
ALA I G C C A 770	ILE VATTG' 830	LYS 1 A A A G 890	GLN ( CAAG 950	GLY P G G T G 1010	ALA I G C C A 1070
PHE TTTC	PRO C C T )	ASP GAT1	leu TTA	GLU GAA	TYR T A T (
GLU GAA1	VAL GLU PRO ILE VAL ASN THR TTGAACCTATTGTCAATACG 820	GIN ILE ASN GLU ASP ASP LYS ALA LYS LEU AL? CAAATCAATGAAGATGATAAAGCAAAATTGGC 870	SER THR LEU GTACATTA( 940	LYS PHE GLU AATTTGAA 1000	GLY G G C 1
ASN A A T C 760	ASN ALA THR VAL A A C G C A A C A G T T 810 820	GLU G A A G 880	SER A G T 940	LYS A A A 1000	ASN A A T 1060
LEU TTA	THR ACA	ASN A A T C	SER AGC A	ASN AAT 1	ASN A A T
LYS LEU AAATT?	ASN ALA THR NACGCAACA 110	GIN ILE PAAATCI	VAL SER 3 T T A G T <i>1</i> 330	TRP GLY 1 G G G G A 1 190	LEU TTA
GLU G A A 750	ASN A A C 810		VAL G T T 930	TRP T G G 990	TYR 1 A T 1050
ARG C G A (	TYR T A T	LEU ILE TTAATT	SER TCTG	LEU TTA	TYR T A T
ILE A T T	ARG C G C	LEU TTA	G A A	LYS A A A ?	ASP G A T
LEU TTTAP 740	. GLY А G G T 800	) ILE AATT 860	G A A C (	2 TRP G T G G A 980	ARG CGTG
WAL G T T	VAL GTA	CLU GAA 8	6 6 6 6	TRP TGG	ILE AATTC 1040
ASP G A T	SER A G T	ASN ARG ALA A A T C G C G C T ( 850	THR PHE LYS CTTTCAAGO	ASP SER 3 A T T C T ?	LEU GIN SER PTGCAGTCAA 1030
G C C	ALA G C A I	ARG C G C	発医 T T C	ASP G A T	GLN C A G
VAL G T T C 730	TYR T A T G 790	ASN A A T 850	THR A C T 910	PRO C C T G 970	LEU T T G 1030
LYS A A A C	HIS C A C T	ASN A A T A	LEU TTA 1	LEU GLN TACAAC	ASP GAT1
PHE LXS VAL GLY ASP VAL LEU ILE ARG GLU LYS LEU ASN GLU PHE ALA LYS SER GTTTAAAGTTGGCGATGTTTTAATTCGAGAAAATTAAATGAATTTGCCAAAAGT 730 770	GLU HIS TYR ALA SER VAL GLY ARG TYR A G A G C A C T A T G C A A G T G T A G G T C G C T A T 790 800	PRO ASN ASN ARG ALA GLU ACCAAATAATCGCGCTGAAA 850	SER LEU THR PHE LYS GLY ASN GLU SER VAL SER SER THR LEU GLN GLU GLN MET GLA A T C A T T A A C T T T C A A G G G G A A C G A A T C T G T A G T A C A T T A C A A G A A C A A T G G A A C A A T G G A A C A A T G G A C A A T G G A C A T T A C A T T A C A T G A A C A A T G G A A C A A T G G A C A A T G G A C A A T G G A C A A T G G A C A T T T A C A T T T T	LEU GLA PRO ASP SER TRP TRP LYS LEU TRP GLY ASN LYS PHE GLU GLY ALA GLA PHE GL ATTACAACCTGATTCTTGGAGAATTATGGGGAAATTAAATTTGAAGGTGCGCAATTCGA 970 980 990 1020	LYS ASP LEU GLN SER ILE ARG ASP TYR LEU ASN ASN GLY TYR ALA LYS ALA GLN ILL GAAAGATTTGCAGTCAATTCGTGATTATTTAATTAATGGCTATGCCAAAGCACAAAT 1030 1030 1040 1050 1050

		12/	82		
MET r A T 1200	SER ' A G 1260	SER T A G 1320	LEU C.T 1380	VAL 2 G T 1440	1 SER TTC 1500
GLY G G T	ARG C G T	G G T	THR A C C	THR ACC	ASN SER AATTC 1500
GLY G G A	ARG C G C	TYR r a c	ILE A T A	ASN A A T	TYR A T
LEU CTC	PHE TC	GLY : G T	ALA C G 1	GLY GA A	irp GG 1
ASN CLU CLY LEU CLA TYR ASP LEU ARG SER ALA ARG ILE ILE CLY ASN LEU CLY CLY MET A A A T G A A G G T T T A C A C C T T C G T A G C C A T T A T A G G T A A T C T G G G A G G T A T T A T A G G T A T T A T A	SFR ALA GLU LEU GLU PRO LEU LEU SER ALA LEU HIS LEU ASN ASP THR PHE ARG ARG SET GTCTGCCGAGCTTGAACCTTTACTTTCAGCATTACATTTAAATGATACTTTCCGCCGTAG 1210 1220 1230 1230	ASP ILE ALA ASP VAL GLU ASN ALA ILE LYS ALA LYS LEU GLY GLU ARG GLY TYR GLY SER TGATATTGCAGATGTAGAAAATGCAATTAAAGCAAACTTGGAGAACGGGTTACGGTAG 1270 1220 1320	ALA THR VAL ASN SER VAL PRO ASP PHE ASP ASP ALA ASN LYS THR LEU ALA ILE THR LEU CGCAACGGTAAATTCAGTACCTGATTTTGATGATGCAAATAAAACATTAGCGATAACCT 1330 1340 1350 1360 1360	G LEU THR VAL ARG GIN LEU ARG PHE GLU GLY ASN THR VAL TTTAACTGTTCGCTTTGAAGGAAATACCGT 1410 1420 1430	SER ALA ASP SER THR LEU ARG GIN GLU MET ARG GIN GIN GLU GLY THR TRP TYR ASN SET TTCTGCTGATAGCACTTTACGTCAGGAAATGCGCCAACAAGAAGGAACTTGGTATAATTC 1450 1450 1460 1460 1470 1500
GLY GGT	ASP GAT	aw Gaaa	THR ACA	PRE L T T	GLY 3 G A
ILE A T A	ASN A A T	GLY 3 G A (	LYS A A A A	ARG : G C 7	GLU AAC
1LE A T T 1180	LEU F T A 7 1240	LEU CTT(	ASN A A T 1 1360	LEU 7 T T ( 1420	GLN : A A G 1480
ARG C G C	LEU GLU PRO LEU LEU SER ALA LEU HIS TTGAACCTTTACTTTCAGCATTACATT 1220	ASP VAL GLU ASN ALA ILE LYS ALA LYS 3 A T G T A G A A A A G C A A A C C A A A A C C A A T T A A A G C A A A A A C A A A A A A A C A	ALA 3 C A 1	THR VAL ARG GIN LEU ACTGTTCGCCAACTTC 1410 1420	GIN : A A C
ALA G C A	LEU FTA	ALA .	ASP 3 A T (	ARG G C C	ARG G C C
SER TAGTG 1170	ALA 1 G C A 1	LYS . A A A (	ASP 3 A T ( 350	VAL F G T T C 1410	MET A A T G C 1470
ARG CGTA 11	SER TCA	ILE ATT	FE 1.	THR A C T (	GLU 3 A A A 14
LEU CTT	LEU	ALA G C A	ASN SER VAL PRO ASP NATTCAGTACCTGATT 1340	LEU	GIN
ASP LEU GACCT1 0	LEU TTA 0	ASN A A T	PRO C C T (	ARG C G T C	ARG C G T C
TYR 7 TATG 1160	PRO I C C T T 1220	GLU 3 A A 2 128	VAL 3 T A ( 134(	ARG ARG CGACG1 1400	LEU 7
GLN CAG	GU GAAA	VAL 3 T A (	SER FCA(	GLY 3 G A (	THR 1 C T 1
LEU T'TA	LEU	ASP 3 A T (	ASN A A T	ALA 3 C T C	SER G C A
GLU GLY LEU GLN TYR 3 A A G G T T T A C A G T A T 6 1150	ALA GLU ; C C G A G ( 1210	ILE ALA T T G C A C 1270	ALA THR VAL CAACGGTA1	VAL VAL ASP ALA GLY TTGTTGATGCTGGAC 1390	ASP 3 A T A 1450
CLU GAAA	ALA 3 C C	ILE A T T (	THR A C G (	WAL	ALA ; C T G
ASN A A T (	SER FCTC	ASP 3 A T A	ALA ; C A ?	VAL VAL ASP ALA GLY ARG AR TGTTGTTGATGCTGGACGACG 1390 1400	SER CTG
A	G 1	E E	ຽ	T G	T

SE

SER LYS

TYR ASP ASN

ASIN

3

呂

呂

ASN VAL

GLY

SER LEU GLY

ASP GLY VAL

LYS

1810

TAAAGATGGTGTAAGTCTTGGTGGAAATGTTTTCTTTGAAAACTACGATAACTACTAAAAG

		13/8	32	
GLU G A 1560	LYS A A 1620	T A G 1680	ALA A G C 1740	THR LAC 1800
VAL G T C	VAL G T C	ILE A T T	ILE A T A	PHE TTT 1
THR ACA	LYS A A A	GLY 3 G T	SER A G T	TYR LAT
GLU 3 A A A	TYR	SER G T (	VAL	PRO CCJ
РНЕ ТС (1550	VAL T A T 1610	3LU ST AGA( 1670	ALA VA C A G 1 1730	SLU PR AGCC 1790
PHE PHE GLU THR VAL GLU TCTTCGAAACAGTCGA 1550 1560	AL T	語 CAG	( AT)	) ຄວວ
I T I	SP 1	X J A T S	7 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	R T T
G G	A 5 6	ម ១ ១	B 0 0	TT A
THR A C 1 154(	WAL G T 0 1600	TYR T A C 1660	THR A C A 1720	G G T 1780
ARG C G T	GLU GAA	GLY 3 G T	GLY 3 G A	LEU r T G
ASP G A T	ASP G A T	ILE A T T	LEU FTG(	AGN A A T
LEU T. T. A 530	ASN A A T 590	GLY G G T . 650	P把 LTC: 710	T G T C 1
ARG CGC	SER A G T	PHE LTT(	ASN A A T :	SER AGTO
ILE A T T (	ILE ASN GLY SER ASN ASP GLU VAL ASP VAL VAL TYR LYS VAL LYS \TCAATGGTAGTGATGAAGTGGATGTCGTATATAAGTCAA 1580 1590 1620	ILE ASN PHE GLY ILE GLY TYR CLY THR GLU SER GLY ILE SER A T C A A C T T T G G T T A C G G T A C A G A G T G G T A T T A G 1670 1680	IN ASP ASN PHE LEUGLY THROUY ALA ALA VALSER ILE ALA A A G A T A A T T T C T T G G G A A C A G G G G G G C A G T A A G T A T A G C C I A A G T A T A G C C I A A G T A T A G C C A G T A A G T A T A G C C C A G T A A G T A T A G C A G T A T A G C A G T A T A G C A G T A T A G C A G T A T A G T A T A G T A T A G T A T A	<b>排</b> 「C G A
LYS A A A A	ASN A A T (	ILE \TC1	GIN A A C	GLY 5 G T 7
GIN LEU VAL GLU LEU GLY LYS ILE ARG LEU ASP ARG THR GLY PHE PHE GLU THR VAL GLAATTAGATTAGATTCGTACAGGTTTCTTCGAAACAGTCGAAACAGTCGAAACAGTCGAAACAGTTGATTCTTCGAAACAGTCGAAACAGTTGAAACAGTTGAAACAGTTGAAACAGTCAGAAACAGTCAGAAACAGTCGAAAACAGTCGAAAACAGTCGAAAACAGTCGAAAACAGTCGAAAACAGTCGAAAACAGTCAAAACAGTCAAAAAAAA	ASN ARG ILE ASP PRO ILE ASN GLY SER ASN ASP GLU VAL ASP VAL VAL TYR LYS VAL LYS A A A C C G A A T G A T C A A T G G T A G T G A G T G G A T G T C G T A T A A A G T C A A A G T C A A A G	GLU ARG ASN THR GLY SER ILE ASN PHE GLY ILE GLY TYR GLY THR GLU SER GLY ILE SER A G A A C G T A A C G G G T A T T G G T A T T G G T T A C G G T A C A G A G A G A G T G T A T T A G 1630 1640 1650 1680	TYR GIN ALA SER VAL LYS GIN ASP ASN PHE LEU GLY THR GLY ALA ALA VAL SER ILE ALA TTATCAAGCAAGTGTTAAACAAGATAATTTCTTGGGAACAGGGGGGGG	GLY THR LYS ASN ASP TYR GLY THR SER VAL ASN LEU GLY TYR THR GLU PRO TYR PHE THR TGGTACGAAAATGATTATGGTACGAGTGTCAATTTGGGTTATACGGAGCCCTATTTAC 1750 1750 1760 1770 1770 1780
LEU TTA	PRO C C T A	GEN	VAL 3 T T 2	ASP 3 A T 1
GLU 3 A G	ASP 3 A T C	出 い で ほ	SER GT (	ASN A T (
VAL 3 T T ( 1510	ILE 1 T T C 1570	ASN A C A 1630	ALA ; C A A 1690	LYS 1 A A A 1750
LEU T A C	ARG ILE GAATTG 1570	ARG G T A	an A A G	FR CGA
GLN	ASN 1	GLU 1	w E	'n F A
GL A C A	ASA A A A	GUA A G A	TYI I T A	<u>പ്ര</u> വാ
	~	• •	•	

		14/	82			
PRO C C C 1920	ASN A A L980	• •	<b>.</b> .	SER T C 2160	ASP A G A 2220	LYS 2 A A 2280
PHE T T C C 19	SER AGT	ASN A A T	ARG A G A	GLY G G T	ARG A G A	ASN A A C
GLY 3 G T	ILE A T T	GLY G G T	ASN A A T	PRO C C A	ASP G A C	GLY G G A
LEU TAC	LYS A A A	LYS A A A (	LEU	ILE \ T T (	LEU LTA(	PHE GLY ASN TTGGAAACI
THR LEU ACTTTA 1910	ASN A T A 1970	PHE L TTA	SER LEU ASN AGCCTTAAT1 2090	THR CT 7 2150	PRO LES CATT 2210	GLY PHE G G T T T 2270
AL 'TTA	MR A	YS A A T	ASN A C A	ARG VAL THR ILE PRO GLY CGAGTTACTATTCCAGGT? 2150	PHE TYR PRO LEU ASP ARG PTCTACCCATTAGACAGAC 2210	A T G
SN A	FR J	et i Tga	ASN TYR ASN AACTATAACI 2080	LRG G A G	插 T C T	ALA ASN SCAAATO
T O	T A T	E A O	_ D		T T 0	7 P P P P P P P P P P P P P P P P P P P
SER A	HIS C A 1	SER T C A 2020	ASN 1 A A 2	GLY 1 G G 214	GLY 5 G G T 2200	TYR 1 T A T 2260
GLY G G A	G G T	GIN C A A	TRP TGG	GLY G G 1	GLN C A G	ALA GLY TYR SCAGGATATO 2260
TYR I A C	LEU I T A	ILE A T T	GLY TRP IGGTTGG1	SER LEU GLY GLY GTCTTGGTGGAC	VAL G T A	ALA G C A
ASP THR SER SER ASN TYR LYS ARG THR THR TYR GLY SER ASN VAL THR LEU GLY PHE PRO TGATACATCCTCTAACTATAAGCGTACGACTTACGGAAGTAATGTTACTTTAGGTTICCC 1870 1870 1880 1890 1890	VAL ASN GLU ASN ASN SER TYR TYR VAL GLY LEU GLY HIS THR TYR ASN LYS ILE SER ASN TGTAAATGAAATAACTCCTATTATGTAGGATTAGGTCATACCTATAATAAATTAGTAA 1930 1930 1940 1950 1950	PHE ALA LEU GLU TYR ASN ARG ASN LEU TYR ILE GLN SER MET LYS PHE LYS GLY ASN CLY CTTTGCTCTAGAATATAACCGTAATTTATATATTCAATCAA	ILE LYS THR ASN ASP PHE ASP PHE SER PHE GLY TRP ASN TYR ASN SER LEU ASN ARG GLY CATTAAAACAAATGACTTTGATTTTTTTTTTGGTTGGAACTATAACAGCCTTAATAGAGG 2050 2050 2060 2070	TYR PHE PRO THR LYS GLY VAL LYS ALA SER LEU GLY GLY ARG VAL THR ILE PRO GLY SE CTATTTCCCAACTAAAGGGGTTAAAGCAAGTCTTGGTGGACGAGTTACTATTCCAGGTTC 2110 2120 2130 2130	ASP ASN LYS TYR TYR LYS LEU SER ALA ASP VAL GLN GLY PHE TYR PRO LEU ASP ARG ASI TGATAACAAATACTACAAACTAAGTGCAGATGTACAGGGTTTCTACCCCATTAGACAGAGA 2170 2210 2220	HIS LEU TRP VAL VAL SER ALA LYS ALA SER ALA GLY TYR ALA ASN GLY PHE GLY ASN LY. TCACCTCTGGGTTGTATCTGCAAAAGCATCTGCAGGATATGCAAATGGTTTTGGAAACAA 2230 2230 2230
THR ACG	VAL G T A	LEU TTA	SER TCT	ALA G C A A 21	ALA G C A	ALA G C A 1 22
ARG C G T	TYR	ASN A A T	PHE TTT	LYS A A A G	SER A G T	LYS A A A
LYS A A G 0	TYR TAT 0	ARG C G T 0	ASP GATT 0	VAL GTTA	LEU CTA	ALA G C A A
TYR T A T A 1880	SER TCC 194	ASN A A C C 2000	PHE TTT 206	GLY G G G 212	LYS A A A C 2180	SER TCTG 2240
ASN AACT	ASN GLU ASN ASN AATGAAAATAAC 1930	PHE ALA LEU GLU TYR ASN TTGCTCTAGAATATAAC 1990 20	ASP GAC'	TYR PHE PRO THR LYS CLY ATTTCCCAACTAAAGGG 2110	ASP ASN LYS TYR TYR LYS A T A A C A A A T A C T A C A A 1 2170	HIS LEU TRP VAL VAL ACCTCTGGGTTGTA 2230
SER SER 1 C C T C T 7 1870	ASN A A T	GLU 3 A A	ASN A A T	THR ACT	TYR TAC	VAL G T T
SER r c c ' 1870	GLU G A A 2 1930	LEU C T A (	ILE LYS THR ASN TTAAAACAAATO 2050	PRO 2 C A 2110	LYS A A A 2170	TRP r G G 2230
THR ACAT	ASN A T (	ALA C T C	LYS A A A	PHE TC	ASN A C	LEU T C
ASP '	VAL S T A A	면 - 면 - 면	표 - - 표 -	κ L	F T	S C C
AS G A	VA G T	T T	IL A T	TY T A	AS G A	HI C A
H	Ę		S	S	E <del>-</del>	€

15/82					
ALA	GLY	ALA	SER	LEU	SER
r G C	r G G	: G C	7 T C	A T T	T C
2340	2400	2460	2520	2580	2640
所 エエエ ( 23	THR ACTO	SER AGC	THR ACC	GLY GGA	ALA SER 3.C C T C 2640
ARG GLY	SER GLY	ALA	ARG	ASN	ARG
GTGGT7		3 C T	C G A I	A A T	C G C (
ARG	SER AGT(	THR ALA ACAGCT )	WAL 3 T C C	LYS ASN GLY LEU AAAATGGATT O	ILE TT
HR TYR THR ALA GLY GLY ILE GLY SER LEU ARG GLY PHE ALZ	TYR GLY ASN GLY S	ASN ALA ILE ALA TA A A A TGCAATCGCTA (2440)	SER GIN ASN THR V	ASP I	ARG ILE ARG
CTTATACAGCGGGGGATCGGTTCATTACGTGGTTTGC	FATGGTAATGGTA		VGCCAAAATACGG	G A T A	CGTATTCGCC
2310 2320 2330	2380		2500	2570	2630
SER	ASN	ILE	ASN	LYS SER	SER
	A A T	A T C	A A T	AATCA(	AGCC
G G T T	GLY	ALA	GLN	LYS	SER
	G G T	G C A	C A A	A A A	TCAA
ILE	TYR	ASN	SER	TRP	LYS
A T C (	T A T	A A T	A G C	F G G A	A A A T
2320	2380	2440	2500	2560	2620
CLY CLY	ALA GLU	GLY 3 G T	LYS A A G I	LYS A A A T	GLY 3 G C 1
G G T	ALA G C C	ILE GLY GLY \TTGGTGGT1	ASP GATP	TRP ASN THR GGAATACTA 2550	TYR I A T (
ALA	TYR	ILE	VAL SER	ASN	PRO ASP TYR
1 G C G (	F T A T C	3 A T T	GTGAGCG	3 A A T	CTGATTAT
2310	2370	2430	2490	2550	2610
THR ACAC 23	ILE ATT	VAL GTG1	WAL GTG 2	TRP TGG	PRO C C T (
TYR T A T	AIA G C A A	SER ASP CTGAT )	PHE ATTT	VAL GTT?	LEU
THR ACT 0	ASN A A C 0	SER T C T	PRO C C A O	SER AGTO	ARG A G.A 7
PHE TYR GLN THR TYR	PRO A	SER S	PRO THR ECA CT C 2480	PHE VAL ASP ALA ALA SER	ASP VAL LEU LYS ARG
TCTATCAAACTTATA	C C T A	AGTT		TTGTTGATGCGGCAAGT	3 A T G T A T T A A A A G A
2300	2360	2420		2530 2540	2590
TYR	ILE GLY	LYS ILE	PRO	ALA	LEU
T A T	TTGGA	AGATA	C C A	3 C G (	FTA
PHE	ILE	LYS	WAL	ASP	VAL
TTC	A T T	A A G	3 T G C	3 A T	3 T A 3
PRO	SER	LYS	ILE	VAL	ASP
C C G T	A G T A	A A G 2	A T T (	3 T T (	3 A T (
2290	2350	2410	2470	2530	2590
LEU TTA	TYR GLY 'ATGGT!	THR PHE LYS CTTTTAAGA	LEU FTA?	PE [ T T (	SER AGCG
ARG LEU PRO PHE TYR GLN TI G C G T T T A C C G T T C T A T C A A A ( 2300	TYR GLY SER ILE GLY PRO ASN ALA ILE TYR ALA GLU TYR GLY ASN GLY SER GLY THR GL) TTATGGTAGTATTGGACCTAACGCAATTTATGCCGAATATGGTAATGGTAGTGGTACTGG 2350 2350 2360 2360 2370	THR PHE LYS LYS ILE SER SER ASP VAL ILE CLY CLY ASN ALA ILE ALA THR ALA SER ALA TACTTTTAAGAAGATAAGTTCTGATGTGATTGGTGGTAATGCAATCGCTACAGCTAGCGC 2410 2420 2430 2430 2460	GLU LEU ILE VAL PRO THR PRO PHE VAL SER ASP LYS SER GLN ASN THR VAL ARG THR SET A G A G T T A A T T G C C A T T T G T G A G C G A G A G C C A A A A T A C G G T C C G A A C C T C 2470 2480 2520	LEU PHE VAL ASP ALA ALA SER VAL TRP ASN THR LYS TRP LYS SER ASP LYS ASN GLY LET CTTATTTGTTGATGCGCAAGTGTTTGGAATACTAAATGGAAATGGATAAAAATGGATT 2530 2540 2580	GLU SER ASP VAL LEU LYS ARG LEU PRO ASP TYR GLY LYS SER SER ARG ILE ARG ALA SEA A G A G A G C G A T T A A A A A A A T T G C C T G A T T A T G G C A A A T C A A G C C G T A T T C G C G C C T C 2590 2690 2600 2610 2610

LYS	A A 2700	SER T C 2760
ALA	TGGCAATCTCCTATTGGGCCATTGGTATTCTCTTATGCCAA	ASN ASP ASP VAL GLU GIN PHE GIN PHE SER ILE GLY GLY SEL A A T G A T G T C C A A T T T A G T T C G A G G T T C C A A T T T A G T T G G A G G T T C C A A T T T A G T T G G A G G T T C C A A T T T A G T T G G A G G T T C C A A T T T A G T T G G A G G T T C C A A T T T A G T T G G A G G T T C C A A T T T A G T T G G A G G T T C C A A T T T A G T A T G G A G G T T C C A A T T T A G T A T G G A G G T T C C A A T T T A G T A T G G A G G T T C C A A T T T A G T A T G G A G G T T C C A A T T T A G T A T T G G A G G T T C C A A T T T A G T A T T G G A G G T T C C A A T T T A G T A T G G A G G T T C C A A T T T A G T A T T T G G A G G T T C C A A T T T T A G T T T G T T T T T T T T
TYR	T A T	G G A
SER	T C T	ILE ATT 50
出	ттст 2690	SER A G T A 2750
VAL	G T A	PHE TTT
TEN	T T G	GIN C A A
PRO	C C A 2680	PHE T T C 2740
GLY	១ ១ ១	GIN C A G
ILE	ATT	GW GAA
PRO	ССТ 2670	VAL G T C 2730
SER	T C T	ASP G A T
GIN	CAA	ASP GAT
TRP	т с 30	ASN A A T
GIN	C A A 266	GU GAA 2772
盟	T T C	TYR
GLY	G G A	LYS A A A
VAL	G T C 2650	LYS A A A 2710
GLY	G G T	ILE A T T
当	TACAGGTGTCGGATTCCAA 2650 26	PRO ILE LYS LYS TYR GLA ACCAATTAAAAAATATGA 2710

16/82	
2820	A A A T G A A A 2880
2810	ATATTTATC 2870
2800	ттааттаа G G 2860
	ГТАААСССАТСАТТТААТТАА G G A T A T T T A T C A A A T G A A A C 2850 2850 2860
	A A A T A '
2770 2780 2780	TTAATTGGGCAGAGA. 2830

CTTCATCAGAACTCAAAAACAACGTTCTCTGCCTAAT

ບ ບ	2940
'TGCTTTAGGTATTGCACTTGCTTCAGGCTATGCTTC	Ö
₽	
ပ	
ග	
E	
A	
Ľ	
ပ	
G	2930
9	23
<b>~</b> ;	
C )	
<u>.</u>	
- [→	
ت	0
Z	2920
ن	2
Ö	
E	
E	
Ø	
E	
<u>ෆ</u>	
G	
<b>~</b>	2910
_	23
<u>-</u>	
[ ]	
(2)	
- [→	
CGCACT	
Ø	
ပ	2
G	2900
ပ	• •
ပ	
×	
⋖	
E	
<u>ن</u> س	
~	
~	0
~	2890
ິບ	Ñ
ပ်	
C	
₽	
æ	
ပ	
¥	
V	

FIG. 1C

DS-691-1-5 DNA, Mirn A D15 sequence IS THE SEQUENCE BEING TRANSLATED

	17	82		
T T G C A A A A G	C A C T G A T T A	G T C A T T T A G	C A A A G C A T C T	ттаат сатт
	120	180	240	300
G T G G C C A A T T T C T A T T G	TTGGATTGGTATTTTTTAAGTTTTATGGCA	NTTAGATGGCGGT	C G G G T G	GTTTGCATTATTAA
40 50	100 110	170		290
A A T A A T T T A A G T G G G	G T A T T T T T T A	CATTACCAGTATT	ААССТGТТТСТGАG	GCTTAACGGTG
30 40		160	220	280
САТТА	. T T G G A T T G G T	A T T T A T T T C 150	G T T A A A G G A A A	T G T T A T T A A
G G C G A T T T G T 20	GCACATCAGCAAATA 70 80	A G G G A T T A T G A 140		G G C G C A G C A C
AATCACTTACT	G T G C T G G C A C A	G T G T A A A T T T A	ттттттаасаат 6 6 аа 6 Ст	G T T A T C G A A T T G G C G
10		130	190	250

			18/82			
	360	VAL G T G 420	ALA GLY GCCGGT 480	ARG C G A 540	ALA LYS G C T A A A 600	LYS A A A 660
SER	<b>-</b> 5	ARG C G T	ALA G C C	G G T	ALA G C T	LEUCTT
LEU ILE ALA SER LEU CTBATCCCAACTTA	٠ ر و		ARG C G T	SER AGT	VAL G T G	S ILE LYS GLY ASN SER VAL ILE PRO THR GLU ALA LEU LYS A A T C A A A G G T A A C T C T G T T A T T C C C A C T G A A G C A C T T A A A 630 640 650
ILE	350	LYS ASP ILE AAAGATATT 410	ARG ALA SER LEU PRO VAL CGAGCAAGTTTACCTGTT 460 470	LEÙ PHE VAL SER TTATTCGTAAGT 530	LEU VAL VAL SER VAL CTTGTTGTTAGCGTT 580 590	G. A. A. 650
	<b>4</b> ر	LYS A A A	PRO C C T	PHE TTC	SER AGC	THR ACT
	- - -	ALA G C A	LEU TTA	LEÙ TTA	VAL G T T	PRO C C C 1
MET LYS LEU THE CAPASAN CAPASAN CHEU	# 0	VAI T S	SER AGT	ARG SER CGCTCT	VAL G T T	ILE ATTC
LYS	340	PHE \ TTTG 400	ALA S GCAA 460	ARG : C G C T 520	LEU V CTTG 580	VAL GTTA 640
MET AT C. 2	5 -	PRO C C T	ARG C G A	VAL GTC	VAL GTG	SER T C T
ر ر د	פ ט	ALA G C A	ILE A T C	ILE A T T	ASP G A T	ASN A A C
ا ا	330	ALA G C C (	GIN C A A 4	ASN A A T 510	GLY G G C 570	GLY G G T 630
A T. A	<b>5</b> <b>−</b>	PHETT	GIN GIN ILE CAACAAATC 450	ALA G C T	GU GLY GAAGGC 570	LYS A A A
ن ن د		VAL G T G	J GLU AGAA	WAL GTG	GIN TCAA	ILE A T C
		~ E		_		LYS A A A 620
- E-	<b>c</b> <b>-</b>	THR THR THR THR ACGACAACGAC 380	ASP G A C	ASN AATG	ALA G C G C	ASP VAL SATGTT?
ا ا	- - 	THR ACA	G G T C	ASP G A C	VAL LYS STGAAA(	ASP G A T
E-		THR ACG 0	GIN C A A (	ARG VAL THR ASP GGTGTGACTGACA 490	VAL GTG 30	SER TCAG
ر 1-	310	GLY 1 GGTA 370	VAL (GITC	VAL GTG 49	ASP V G A T G 550	11.E 8 ATTT 610
ر ح	ر •	PHETC	G G T	ARG C G T	ASP GAT(	SER ILE SER ASP VAL LYS TCGATCATTTCAGATGTTAA 610 620
4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	- -	LEUTTA	ASP GLY VAL GLN GLY ASP LEI GATGGTGTCAAGGTGACTT 430	GIN ARG VAL THR ASP ASN ASE CAGCGTGTGACTGACAATGA 490 500	PHE ASP ASP VAL LYS ALA HES TTCGATGATGTGAAAGCGCA 550 560	SER T C G A

		19/82			
J ASN AAAT 720	VAL GLY ARG TYR ASN ALA THR VAL GTAGGTCGCTATAACGCAACAGTT 760 770 780	GLU PRO ILE VAL ASN THR LEU PRO ASN ASN ARG ALA GLU ILE LEU ILE GIN ILE ASN GLU GAACCTATTGTCAATACGCTACCAAATAATCGCGCTGAAATTTAATTCAAATGAA 790 820 830 840	SER AGT 900	GLY ASN LYS GGAAATAAA 960	U LYS ASP LEU GLN SER ILE ARG ASP TYR TYR LEU ASN ASN GAAAGATTTGCAGTCAATTCGTGATTATTATTAAATAAT 990 1020
LEU TTA	THR ACA(	ASN A A T	SER A G C	ASN A A T	ASN A A T
LYS A A A	ALA G C A	ILE A T C	SER A G T	TRP GLY ! G G G G A !50	LEU TTA
ARG GLU LYS CGAGAAAAA 710	ASN A A C G 770	GLN C A A 830	VAL G T T 890	0	TYR 1 T A T
	TYR TAT	ILE A T T	SER T C T	LEU T T A	TYR T A T
ILE A T T	ARG CGC1	LEU TTA	GW GAA	LYS A A A	ASP G A T
VAL LEU ; T T T T A A	, GLY AGGTC 760	) ILE A A T T 820	G A A C (	' TRP G T G G A 940	ARG CGT 00
0	VAL G T A	GLU GAA 82	GLY G G G <i>I</i>	TRE TG	11.E A A T T C 1000
ASP GAT(	ALA SER GCAAGT	ALA G C T	LYS A A G	ASP SER GATTCT	SER T C A A
GLY ASP	ALA G C A	ARG C G C	PHE TTC	ASP GAT	GLN C A G
	TYR T A T 750	ASN A A T 810	THR A C T 870	PRO C C T 930	LEU T T G 990
PHE LYS VAL TTTAAAGTT 690	HIS CAC	ASN A A T	LEU TTA	GIN C A A	ASP G A T
PHETTT	GLU GAG	PRO C C A	SER T C A	LEU TTA	LYS A A A
ASN GLY 1 A C G G G 680	S	LEU C T A 800	ALA G C A 860	n A	GU GAG 980
ASN A A C	VAL G T A	THR ACG	LEU TTG	MET A T G	PE TTC
ALA G C T i	PHE ALA LYS SER VAL LYS TTGCCAAAAGTGTAAA 730	PRO ILE VAL ASN THR LEU CTATTGTCAATACGCT1 790 800	ALA LYS LEU CAAAATTGO	GIN C A A	GLN C A A
J ASP A G A T G 670	A LYS C A A A 730	3 VAL TGTC 790	; ALA A G C A 850	и спл А G A A 910	r ALA TGCG 970
LEU TTAG	ALA G C C	ILE A T T	ASP LYS ; A T A A A (	GIN C A A 9)	G G T
GIN ASN LEU ASP ALA ASN GL CAAAACTTAGATGCTAACGG 670 680		PRO C C T		THR LEU GLN GLU GLN MET GLU ACATTACAAGAACAAATGGA1 910 920	PHE GLU GLY ALA GLN PHE GLU TTTGAAGGTGCGCAATTCGA(
G A A	GLU GAA1	G A A C	ASP G A T (	THR A C A	PAE TTT

		20/82	?		
LYS	ILE A T T T 1140	LEU	LEU	ASN	LEU
A A A		T T A	CTT	A A T	CTT
1080		1200	1260	1320	1380
THR A C A	ARG ILE CGCATT 1140	HIS C A T	ALA LYS ; CAAAA	ALA G C A	GIN C A A
THR LYS THR ASP VAL GIN LEU ASN ASP GLU LYS THR LYS ACTAAAACGGATGTTCAGCTAAATGATGAAAAAAAAAAA	ALA 3 C A	SER ALA GLU LEU GLU PRO LEU LEU SER ALA LEU HIS LEU TCTGCCGAGCTTGAACCTTTACTTCAGCATTACATTA 1170 1180 1190	ALA G C A	VAL ASN SER VAL PRO ASP PHE ASP ASP ALA ASN GTA A A TTCAGTA CCTGATTTTGATGA A A T 1290 1300 1310	ARG C G C
GLU 1 G A A A 1070	3 SER TAGT(	LEU SER ALA TTTCAGCA7	ILE LYS ATTAAA(	РНЕ ASP ТТСАТ( 1310	THR VAL ACTGTT 1370
ASP	ARG	SER	ILE	PHE	THRACT
GATG	CGT	T C A	A T T	TTT	
ASN A A T G	LEU	LEU	ALA G C A	ASP GATT	LEU
LEU	ASP	LEU	ASN	PRO	ARG
CTA	GACO	TTA	AAT	C C T	C G T 1
VAL GIN LEU	GLU GLY LEU GIN TYR ASP LEU	GLU PRO LEU	GLU ASN	SER VAL PRO	ARG 7
STTCAGCTA1	GAAGGTTTACAGTATGACCTT	1 G A A C C T T T A C	GAAAAT	CAGTACCT	C G A C
1060	1110 1120	1180	1240	1300	1360
VAL	GIN	GA A	VAL	SER	GLY
GTT	C A G		G T A	T C A	G G A
ASP	LEU	LEU	ASP	ASN	ALA
GATG	TTA	CTT	GAT(	A A T 1	G C T
THR LYS THR ACTAAAACGO	G G T 1110	GLU LEU GAGCT 1170	ALA G C A 1230	VAL G T A 1 1290	ASP G A T 1350
LYS	GLU	ALA	ILE	ALA THR	VAL
A A A	GAA	G C C	A T T	GCAACG	G T T
THR A C T	ASN	SER	ASP	ALA	VAL
	A A T	T.C.T.C	GAT	G C A I	G T T
1LE	ASP VAL	MET	SER	SER	LEU
1 A T T	3 A T G T A	'A T G	r A G T	1 A G C	; C T T
1040	1100	1160	1220	1280	1340
ALA GIN ILE SCACAAAT 1040	ASP G A T	GLY GLY MET 3 G A G G T A T (	ARG CGT	G G T	THR ACC
ALA	ILE	G G A	ARG	GLU ARG GLY TYR GLY SER	ILE
G C A	ATT G		C G C C	3 A A C G C G G T A C G G T A G (	A T A
ALA LYS S C C A A A G 1030	VAL THR GTAACCI	GLY ASN LEU GTAATCTGO 1150	ASP THR PHE (ATAC)	G G T 70	ALA G C G
ALA I	VAL 7	ASN 1	THR FACTT 1210	ARG C	LEU 7
G C C A	G T A A	AATC		C G C G	TTAG
1030	1090	1150		1270	1330
GLY TYR ALA LYS ALA GLN ILE GGCTATGCCAAAGCACAAAT 1030 1040	ASN A A T		rn		LYS THR LEU ALA ILE THR LEU VAL VAL ASP ALA GLY ARG ARG LEU THR VAL ARG GIN LEU A A A A C A T T A G C G A T A A C C C T T G T T G A T G C T G G A C G T T T A A C T C G C C A A C T T 1330 1330 1340 1350 1350
G G C I	VAL	ILE	ASN	GLY	LYS
	G T T	A T A C	AAT(	GGA	A A A

		21/82			
GLN C A A 1440	THR A C A 1500	VAL G T G 1560	TYR T A C 1620	THR A C A 1680	GLY G G T 1740
ARG GLN GLN: GCCAACAA	ARG THR CGTACA 1500	GLU VAL GAAGTG 1560	GLY GGT	GLY THR GGAACA 1680	LEU TTG
	ASP GAT(	ASP GATG	ILE A T T	PHE LEU GLY FTCTTGGGA 570	ASN A A T
MET A T G (	LEU CTTA (	ASN 'AATC 1550	PHE GLY TTGGT1	PHE 7 T T C 1670	SER VAL ASN LEU GEGECAATTE 1730
GIN GIU CAGGAA1	ARG CGC	SER AGTA 15	PHE TTT	ASN A A T	SER AGT
GINGLUMET ARG CAGGAAATGCGC 1430	SER GIN LEU VAL GLU LEU GLY LYS ILE ARG LEU ASP CACAATTAGTTGAGTTAGGAAAATTCGCTTAGAT 50 1470 1480 1490	ARG ILE ASP PRO ILE ASN GLY SER ASN ASP CGAATTGATCCTATCAATGGTAGTAATGAT 1530 1550	ASN A A C T	SER TYR GIN ALA SER VAL LYS GIN ASP ASN PHE LEU GTTATCAAGCAAGTGTTAAACAAGATAATTTCTTG 10 1650 1670	THR A C G A
ARG C G T C O	LYS A A A A 0	ASN A A T 0	ILE ATCA 0	GLN C A A 0	ASP TYR GLY 3 A T T A T G G T A
LEU P TTAC 1420	LEU GLY L FTAGGAA 1480	ILE AATCA	SER ] A G T A 1600	LYS GAAAC	TYR ( TATG 1720
ASP SER THR LEU ARGGATAGCACTTTACGT	LEU TTA	PRO C C T A	ASN THR GLY SER AACACGGGTAGTA 1590 1600	VAL G T T	ASP G A T
SER A G C	VAL GLU 5 T T G A G 1 1470	ASP G A T C	THR A C G	SER AGTG	ASN A A T G
ASP G A T 1410	VAL G T T 1470	ILE A T T 1530	ASN A A C 1590	ALA G C A 1650	LYS A A A A 1710
ALA G C T	LEU TTA	ARG C G A	ARG C G T	GIN C A A	THR ACG
SER T C T	GIN C A A T	ASN A A C	GLU GAA(	TYR T A T	G G T 1
VAL 3 G T T 1400	SER TCAC 1460	THR VAL CLU ICAGTCGAA 1520	LYS VAL LYS AAAGTCAAA( 1580	SER 7 A G T	ALA 1 G C T (
ACC	ASN A A T	WAL GTC	VAL GTC	GLY ILE 3 G T A T T	ILE ATA 1
ASN A A T A	TYR TAT?	THR ACA	LYS A A A	G G T	SER AGTA
GLY G G A 30	TRP T G G 1 50	не αш тсваа 1510	TYR T A T 70	SER AGT(	ALA ALA VAL SCGCAGTA/ 1690
GLU C G A A G 1390	GLY THR 1 5 G A A C T T 1450	ЭНЕ ТТС 151	VAL 1 G T A T 1570	THR GLU S CAGAGA 1630	ALA (G C A G 1690
ARG PHE GLU GLY ASN THR VAL SER ALA CGCTTTGAAGGAAATACCGTTTCTGCT 1390 1400		GLY PHE PHE GLU THR VAL GLU ASN GGTTTCTTCGAAACAGTCGAAAAC 1510 1520	ASP VAL VAL TYR LYS VAL LYS GLU ARG ASN THR GLY SER II.E ASN PHE GLY II.E GLY TYR GATGTCGTATATAAAGTCAAGGTACGTAGTAGTATTAGGTAC 1570 1580 1590 1600 1600	FQL:	( - )
ARG C G C	GLU GAA(	GLY G G T	ASP G A T	G G T	C C C C
		`			•

GLU 3 A A 1800	SER A G T 1860	HIS 75 A T 2 S 1920	SER C A 1980	ASN A A C 2040	GLY G G A 2100
PHE GLU TTTGAA 1800	GLY S GGAA	ASN ASN SER TYR TYR VAL GLY LEU GLY HIS A A T A A C T C C T A T T A G G T C A T 1900 1920	GIN SER CAATCA 1980	TRP P TGGA	GLY G
PHE I	TYR (	LEU (		GGTT	LEU (
VAL : G T T T 1790	THR 6	GLY I G G A T 1910	TYR 1 T A T A 1970	PHE (TTTG 2030	SER A G T C 2090
ASN A A T G	THR A C G A 18	VAL GLY 3 T A G G A 1910	LEU FTAT 19	SER ICTI	ALA 3 C A A 20
GLY	ARG	TYR FAT (	ASN A A T' 1	PHE ITT	LYS A A A C
LEU GLY : T T G G T ( 1780	LYS A A G (	SER TYR CCTATO	ARG C G T 1	PHE ASP TTGATT 2020	VAL STT1
LEU ( CTTG 1780	TYR I TATA 1840	SER T	ASN 1 A A C C 1960	PHE 7 TTTG 2020	GLY VAL 3 G G G T T 2080
VAL SER LEU GLY GLY ASN VAL PHE GTAAGTCTTGGTGGAAATGTTTC 1780 1790	N ASP THR SER SER ASN TYR LYS ARG THR THR TGATACATCCTCTAACTATAAGCGTACGACT 1830 1850	ASN AAC1	ALA LEU GLU TYR ASN ARG ASN LEU TYR ILE GCTCTAGAATATAACCGTAATTTATATATT 1950 1970	ASP PHE ASP PHE SER PHE GACTTTGATTTTTCTTT 2020 2030	TYR PHE PRO THR LYS GLY VAL LYS ALA SER LEU GLY TATTTCCCAACTAAAGGGGTTAAAGCAAGTCTTGGT 2070 2080 2090
VAL G T A	SER T C T	ASN A A T	GLU G A A		THR
GLY G G T 1770	SER T C C 1830	G A A 1890	LEU C T A 1950	LYS THR ASN AAAACAAAT 2010	PRO C C A 2070
LYS ASP AAAGAT(	THR ACA	VAL ASN GLU GTAAATGAA 1890	ALA G C T	LYS A A A	PRE TTC
	ASP G A T	VAL G T A	PHETT	ILE A T T	TYR T A T
РНЕ ТНК ТТАСТ 1760	LYS SER AAAGT 1820	РНЕ РКО ТСССТ 1880	SER ASN GTAAC 1940	ASN GLX IATGC 2000	ARG GLY 1 G A G G C 2060
PHETTI	LYS	PHE	SER	ASN 'A A T	ARG A
TYR	SER	GLY A G G T	ILE A A T T	GLY NGGT	ASN
JU PRO AGCCC1 1750	ASP ASN SER SATAACTCT 1810	THR LEU GLY 1 C T T T A G G T T 1870	ASN LYS ILE NATAAATT 1930	PHE LYS GLY TTAAAGGT 1990	SER LEU GCCTTA 2050
C G A G C 1750	C G A 3	TAC:	T A A 7	ATT?	SER CAG(
TYR THR GLU PRO TYR PHE THE TATACCGAGCCCTATTTAC 1750 1760	ASN TYR ASP ASN SER LYS SEB AACTACGATAACTCTAAAAG 1810 1820	ASN VAL THR LEU GLY PHE PRO A A T G T T A C T T T A G G T T T C C C 1870 1880	THR TYR ASN LYS ILE SER ASI ACCTATAATAAATTAGTAA 1930 1940	MET LYS PHE LYS GLY ASN GL) ATGAATTTAAAGGTAATGG 1990 2000	TYR ASN SER LEU ASN ARG GLY TATAACAGCCTTAATAGAGG 2050 2060
TYR T A T	ASI A A	ASN A A T C	'I'H	MET A T G	TYR T A T /

		23/82	2		
GLY G G T 2160	TYR T A T 2220	ILE A T C 2280	TYR T A T 2340	ASN A A T 2400	SER A G C 2460
GIN C A G	ALA GLY TYR ; C A G G A T A T 2220	2 9 9 KID	GU GAA	GLY GLY ASN GGGGTAAT 2400	LYS A A G
VAL G T A C	ALA G C A	GLY G G T (	ALA G C C		ASP G A T
ASP ASN LYS TYR TYR LYS LEU SER ALA ASP VAL GIN GLY GATAACAAATACTACAAACTAAGTGCAGATGTACAGGGT 2130 2140 2150	SER r c r 210	ARG LEU PRO PHE TYR GIN THR TYR THR ALA GLY GLY ILE CGTTTACGTTCTATCAACTTATACAGCGGGTGGCATC 2250 2260 2280	TYR 7 T A T ( 2330	ASP VAL ILE 3 A T G T G A T T (	SER 3 A G C ( 2450
ALA G C A (	ALA G C A ?	THR A C A	ILE ATT	VAL G T G	VAL GTG
SER AGTG	LYS A A A G	TYR	ALA G C A	ASP G A T	PHETT
LEU CTA:	ALA G C A 1 00	THR A C T T 50	ASN A A C 20	SER T C T (	PRO C C A 1
TYR LYS L PACAAAC 2140	VAL SER ALA LYS GTATCTGCAAAA 2200	TYR GIN 1 ATCAAA 2260	PRO A C C T A 2320	LYS LYS ILE SER SER ASP AAGAAGATAAGTTCTGAT 2370 2370	PROTHRE CAACTC 2440
TYR T A C		TYR T A T	G G A	ILE A T A	PRO C C A
TYR TAC1	TRP VAL TGGGTT( 2190	PHE TTC1	ILE A T T	LYS A A G I	VAL G T G (
LYS A A A ' 2130		PRO C C G 7 2250	SER A G T 2310	LYS A A G .2370	ILE A T T C 2430
ASN A A C	LEU CTC	LEU TTA	GLY GGT	PHETT	LEU TTA
ASP GAT1	HIS C A C		TYR T A T	THRACT	GW GAG
SER 7 T C T 2120	ARG ASP IGAGAT 2180	LYS ? A A G 2240	ALA ' G C T 2300	ст. г.с.с.т 2360	ALA : G C A 2420
PRO GLY SER CAGGTTC	ARG A G A	GLY ASN LYS GAAACAA 2240	GLY PHE ALA 3 G T T T T G C '	GLY THR GLY 3 G T A C T G G '	ALA SER ALA CTAGCGC 2420
PRO D	ASP G A C A	GLY	G G T	G G T	ALA
R ILE TATT ( 2110	O LEU : A T T A C 2170	X PHE 3 T T T C 2230	л ARG 'A C G T C 2290	.X SER 3 T A G T C 2350	ILE ALA THR TCGCTACA( 2410
THR'ACTA	PRO C C A T 2170	ASN GLY 1 A T G G T 1 223(	SER LEU P. C.A.T.T.A.( 229(	GLY G G T A 2350	ALA G C T 24
ARG VAL THR ILE PRO GLY SER CGAGTTACTATTCCAGGTTCT 2110	PHE TYR PRO LEU ASP ARG ASP TTCTACCCATTAGACAGAT 2170 2180	<b>F</b>	<b>—</b>	GLY ASN GLY SER GLY THR GLY GGTAATGGTAGTGGT 2350 2360	
ARG C G A (	PHE TTC'	ALA G C A J	GLY G G T ?	G G T	ALA G C A A

			24/82		
TRP	T G G 2520	LYS A A A 2580	GLY PRO GGGCCA 2640	РНЕ ТТС 2700	A A A 2760
LYS	AAA	G G C	G G G	GIN C A G	CTC
混	GTTGATGCGGCAAGTGTTTGGAATACTAAATGG 2490 2520	ASP VAL LEU LYS ARG LEU PRO ASP TYR GLY LYS GATGTATTAAAAGATTGCCTGATTATGCAAA 2550 2550 2570 2580	ILE A T T	GLU 3 A A	PHE *** *** TTCTAATAAATTGAACTTTTTTCTTCAGAACTCAAA 2730 2740 2750 2760
TRP ASN	3 A A T 2510	ASP ' G A T T 2570	PRO C C T 2 2630	VAL 'GTC( 2690	1 T C A 2750
	ក	PRO CCTC	SER TCTO	ASP GATO	T C A
VAL	G T T	LEU TTGC	GIN CAA7	ASP GATO	T C T
SER	A G T	ARG A G A 7	GIN TRP CAATGGO 2620	ASN A A T C	T T T 10
ALA ALA	G C A A 2500	LYS P A A A A 2560	GIN C A A T 2620	GLU 7 GAAA 2680	C T T T 2740
	ອ ວ ອ	LEU T T A	N THR GLY VAL GLY PHE TACAGGTGTCGGATTC 2610	LYS TYR GLU AAATATGAAA 2680	GAA
ASP	G A T	VAL G T A 1	GLY GGA	LYS A A A	ATT
VAL	G T T 2490	ASP G A T 2550	VAL G T C 2610	LYS A A A A 2670	*** T A A 2730
别	TTT	SER AGC	G G T	ILE A T T	*** T A A
ren T	ТТА	J GLU SER AGAGAGC	THR ACA	PRO C C A A	雅 TTC'
SER	CTCC 2480	LEU 1 T T A 2540	ALA SER ; C C T C T 2600	TYR ALA LYS PATGCCAAA 2660	SER 1 T C T 2720
TH	ວ ∨ ເ	ASN GLY IATGGAT	ALA G C C	ALA G C C	GLY SER G G T T C ' 2720
ARG	C G A	ASN A A T	ARG C G C G	TYR T A T	GCAG
THR VAL	G 7 C	LYS AAAA 30	ARG ILE : G T A T T C 2590	SER TCT1	ILE ATT (
当	CAAAATACGGTCCGAACCTCCTTATTT 2470 2480	LYS SER ASP LYS ASN GLY LEGA A A A A T C A G A T A A A A A T G G A T T 2530 2540	SER SER ARG ILE ARG ALA SER TCAAGCCGTATTCGCGCCTC 2590 2600	LEU VAL PHE SER TYR ALA LYS TTGGTATTCTCTTATGCCAA 2650 2660	GIN PHE SER ILE GLY GLY SER CAATTTAGTATTGGAGGTTC 2710 2720
ASIN	A T T	SER T C A (	SER AGCC	VAL GTA1	PHE TTTA
CIN	CA	LYS A A A 7	SER TCA A	LEU TTG(	GIN C A A '

ATTA	2820
CCCATCATTTA	2810
AAAATATAAA	2800
TGGGCAGAG	2790
CTAATTTAAT	2780
AACAACGTTCTCTGC	2770

TATTGCAC	2880	
TGCTTTAGGTA	2870	
TAACCGCACT	2860	
TCGCAAAAG	2850	
ATGAAAACA	2840	
AGGATATTTATCAA	2830	

T A T A N T T	2940
PAATGCGGGTT	2930
T T G C T T T C A T T A A T	2920
GAAGAAAAATTGCTI	2910
GCTTCCGCTG	2900
TGCTTCAGGCTATC	2890

25/82

TNCAAGGCNAAGG 2950

FIG. 1D.

	TRANSLATED
	BEING
	SEQUENCE
SB33 D15	IS THE

		26/92		
TACTT	CTGGC	TAAAT 8/92	TTTTA	A T C G A
60	120		240	300
CTTGAAATATTAGGGAAATTA	TGCAAAAGGTG	A T T G A T T A G T G	тсаттта 6 ттт	CAAAGCATCTGTTATC
40 50	110	170	230	
ССТТ G А А А А Т 40	СААТТТСТАТТ	3 T T T T A T G G C 160	GTATTAGATGGCGG	A G C G G G T G C A 280
тсссттттаас 30	AATTTAAGTGGGCO	G T G T A T T T T T A A G 150	ттасса 210	C C T G T T T C T G 2
AGGACAGCTTT	H	G A T T G	T T T C C A	CTGTTAAAGGAAAA
20		140	200	260
GGCATTGAAAAACAG	ACTGGCGATTTGTCATTAAA	G C A T C A G C A A A T A T T G	ттасссаттатсааттта	АСААТ G G A A G C T G T
10	70 80	130	190	250

		27/8	32		
A C G T 360	GLY G G 420	VAL G T 480	VAL G T 540	ASP G A 600	ILE A T 660
T A C	PHE TTC	GLX G G T	ARG C G T	ASP G A T	ILE A T C
T T T	LEU TTA	ASP G A T	GIN C A G	PHE TTC	SER T C G
E→	Leu T. A	VAL 3 T G (	GLY 3 G T	G A	LYS A A A '
A A T G A T 350	SER ] AGTT 410	ARG VAL CGTGTG 470	AIA ( G C C G 530	GУ 1 GGTC 590	ALA LYS G C T A A A 650
TTA	ALA G C A I	ILE ATT(	ARG C G T (	SER AGT(	WAL GTG(
T A T	ILE A T C (	ASP 3 A T I	VAL G T T (	VAL G T A	VAL, G T T (
G C A T 340	LEU CTA	ALA LYS ASP GCAAAAGAT 460	PRO C C T ( 520	所 T T C ( 580	SER A G C ( 640
9 H	LEU CTT	ALA 3 C A 2	LEU TTA	LEU TTA'	VAL G T T
	LYS A A A (	VAL G T G (	SER AGT	SER	VAL GTT
330 330	LYS A A A A 390	PHE T T T 450	ALA G C A . 510	ARG C G C 5	LEU CTT 630
T A A	MET A T G	PRO C C T	ARG C G A	VAL G T C	VAL G T G
÷ د و	ຽ		ILE A T C	ILE A T T	ASP GAT
A A	CAATC	ALA G C C	GLN C A A 10		G G C
320	A T A C 380	PHE ALA TTTGC(	GIN G CAAC 500	AIA 7 G C T A 560	HIS GIN GIU GLY ATCAAGAAGG(
<b>→</b> 9	A G G	WAL GTG1	LEU CLU PAGAA(	ASP VAL	GIN C A A
ر ¥ ر	T A T	THR ACT (	LEU TTA	ASP G A T	HIS C A T
310	T T A 370	THR A C G 7 430	ASP G A C 1 490	ASN A A T ( 550	ALA G C G 610
כ פ	AAT	THR THR .CGACA?	GIN GLY AAGGTO	ASP G A C	VAL LYS ALA TGAAAGCGO
310	CTATAATTTATAGGATAC 370	THR THR THR VAL PHE ALA ALA TACGACACACTGTTTGCCGCA 430 440	GLN GLY ASP LEU GLU GLN GLN TCAAGGTGACTTAGAACAACAA 490 500	THR ASP ASN ASP VAL ALA ASN GACTGACAATGATGTGGCTAAT 550 560	VAL LYS ALA HIS GIN GLU GLY ASP TGTGAAAGCGCATCAAGAAGGCGAT 610 620

		28/8	12		
LEU T. T 720	ALA G C 780	ILE A T 6	LYS A A 900	GLN C A 960	GLY 1 G G 1020
ASN LEI AACTT 720	PHE I T T	PRO C C T	ASP G A T	LEU GL TTACA 960	GLU ; A A C
GLN C A A	GLU PHE AL GAATITGC 780	GLU GAA(	ASP 3 A T (	THR 1 C A 1	PHE 'TTG
LYS	ASN	VAL	GLU	SER	LYS
A A A (	1 A T (	3 T T (	3 A A C	\ G T A	A A A T
LEU LYS CTAAAA 710	LEU L LTAA 770	THR 1	ASN GLU ASP AATGAAGAT 890	SER SER SER THR AGTAGCAGTACA 950	ASN 1
ALA	LYS	AIA	IIE	SER	GLY
G C A (	\ A A 7	3 C A 7	TC?	1 G T A	3 G A A
GLU	GLU	ASN	GLN ILE	VAL	TRP GGG
GAA(	3 A A A	1 A C G	CAAATC	G T T A	
PRO	LEU ILE ARG GLU LYS LEU ASN	VAL GLY ARG TYR ASN ALA THR VAL	ILE	SER	TRP LYS LEU TRP GLY ASN LYS PHE GLU GL)
C C T	TTAATTCGAGAAAAATTAAAT	GTAGGTCGCTATAACGCAACCGTT	ATTC	TCTG	TGGAAATTATGGGGAAATAAATTTGAAGG
700	760	810 820 830	880	940	1000 1020
PRO	ILE	ARG	LEU	GLU	LYS
C C A	A T T (	C G C	FTA/	GAA7	A A A 7
ile a t t	LEU	GGT (	ILE A T T	ASN AAC(	TRP I G G A
ILE	11.E	VAL	GU ILE LEU ILE	GLY	SER TRP
A T T	A T T	G T A 0	GAAATTTTAATT	G G G 7	CTTGG1
690	750	810	870 880	930	990
IAACTCT	ASP G A T	ALA SER GCAAGT	ALA G C T	LYS A A G	V, E.,
ASN	2 9 9	ALA	ARG	PE	ASP G A T 1
A A C	XID	G C A	C G C	TTC	
	LYS VAL AAGTT 740	۳ F	<del>&gt;</del> ₽	~ [-	~ <b>₽</b>
LYS GLY AAAGG9	LYS A A A 74	HIS TYR CACTA1 800	ASN ASN AATAA1 860	LEU THR TTAACT 920	GLN 1 C A A C 980
ILE	PHE	GU	PRO	SER	LEU
A T C A	T T T	GAGC	C C A 1	T C A	FTA(
LYS A A A A	G. G. G. T	VAL LYS TAAAAC 790	THR LEU CCCTAC 850	ALA G C A '	MET GLU LEU TGGAATTA( 970
VAL	ASN	VAL	THR	LEU	MET
G T T 1	A A C (	G T A	A C G	T T G G	A T G (
670	730	790	850	910	970
ASP	ALA	Ser	ASN	ALA LYS	GIN
G A T G	G C T.	A G T C	A A T A	CCAAA1	
SER ASP VAL LYS ILE LYS GL	ASP ALA ASN GLY PHE LYS VA	GIN SER VAL LYS GLU HLS TY	VAL ASN THR LEU PRO ASN AST	ALA LYS LEU ALA SER LEU THE	GLU GIN MET GLU LEU GIN PRC
TTCAGATGTTAAAATCAAAG	AGATGCTAACGGGTTTAAAGT	CCAAAGTGTAAAAGAGCACTA	TGTCAATACGCTACCAAATAA	AGCCAAATTGGCATCATTAAC	AGAACAATGGAATTACAACC
670 680	730	790 800	850 850	910 920	970 980
H	K	ິບ	Ţ	A (	A G

			29/82			
ALA	1080	VAL G T 1140	ASN 1 A A 1200	THR ' A C 1260	ARG 1 C G 1320	LEU . T. T 1380
TYR	r a T	ASN A A T G	GLY G G T	ASP GATA 12	GLU GAA(	THR LEU ACATT 1380
GLY	ວິວ	VAL 3 T T	ILE A T A (	ASN A A T	GLY 3 G G	ASN LYS ATTAAA
ASN GLY	ATC	LXS A A A G	ILE \TT!	LEU FTAA O	LEU GLY	ASN A T A T
ASN	A T A 1070	THR I ACAA 1130	ARG ILE SGCATTA 1190	HIS LEG CATTT 1250	LYS I A A C 1310	ALA P 3 C A A 1370
TEN	TAP	LYS A A A A	ALA G C A C	LEU TAC	ALA LYS 3 C A A A A C 131(	ASP 3 A T G
IYR	ATT	GLU AAA	SER A G T G	ALA C A T	LYS A A A G	ASP 3 A C G
TYR TYR LEU	7 A T T 1060	ASP GLU ; A T G A A 1120	ARG C G T A 1180	SER ALA LEU CAGCATTAC 1240	ILE \ T T A 1300	РИЕ ТТ Т G 1360
ASP	ATT	ASN A T G	LEU	LEU	ALA 3 C A A	ASP 3 A T T
ARG	CGTGATTATTTAAATAATGGCTATGC 1060 1060 1080	LEU		LEU LTAC	ASN A A T G	
ILE		GIN ' C A G C 1110	TYR ASP 5 T A T G A C (	LEU GLU PRO LEU LEU SER ALA LEU HIS LEU ASN ASP THE CTTGAACCTTTACTTCAGCATTACATTAAATGATAC 1230 1240 1260	GLU 1 G A A 1 1290	SER VAL PRO CTGTACCT 1350
ALA	3 C A 1	VAL 3 T T (	GIN TYR CAGTAT 1170	GLU 3 A A C 12	VAL 3 T A C 12	SER FCT(
CIN	CAGGCAATT 1050	ASP 3 A T (	LEU L'T A	LEU CTT(	ASP G A T (	AL ASN SER VAL TAAATTCTGTA 1350
TEN.		ALA G C G (	GLY G T	GLU A G	ALA G C A O	VAL G T A A
ASP	G A T T 1040	LYS A A A A A B A A B A B A B A B A B	GU (GAAG	ALA ( G C C G 1220	ILE A T T G 1280	THR A C A G 1340
LYS	AAA	ALA GIN ILE THR 3 CACAAATCACT 1 1090	ASN A A T C	SER	ASP GATA	THR ACA
PHE GLU	G A G	ILE A T C	ASP VAL 3 A T G T A 1 1150	MET A T G	SER AGTG	ASN A A C
別田	T T C 1030	GLN C A A 1090	ASP G A T 1150	GLY GLY MET 3GAGGTATG1 1210	ARG C G T A 1270	GLY G G G T 1330
GEN	CAA	ALA G C A	ILE A T T (	GLY GGA	ARG C G C (	GLY TYR GLY ASN THR THR GTTACGGTAACACAACA 1330 13
ALA	TGCGCAATTCGAGAAAGATTTG 1030 1040	LYS ALA GIN ILE THR LYS ALA ASP VAL GIN LEU ASN ASP GLU LYS THR LYS VAL ASN VAL CAAAGCACAAAACTAAAGCGGATGTTCAGCTAAATGATGAAAAAAAA	THR ILE ASP VAL ASN CLU A A C C A T G G T G T A A A T G A A G 1150	LEU CLY CLY MET SER ALA TCTGGGAGGTATGTCTGCCG 1210	PHE ARG ARG SER ASP ILE ALA ASP VAL GLU ASN ALA ILE LYS ALA LYS LEU GLY GLU ARC TTTCCGCCGTAGTATTGCAGATGTAGAAATGCAATTAAAGCAAAACTTGGGGAACG 1270 1280 1280 1280 1380	GLY TYR GLY ASN THR THR VAGGTTACGGTAACAG 1330
	₽		Ø	E+	E+	<b>V</b>
		•			•	

		30/82			
GLU	THR	PHE	VAL	THR GLU	ALA
. G A	A A C	7 T T	G T	1 C A G A	; G C
1440	1500	1560	1620	1680	1740
PHE TTT	GLY 3 G A	所E T T C 1	VAL G T C	THR A C A	ALA G C G G 17
ARG	GLU	GLY	ASP	GLY	G G G
C G C	GAA	GGT	G A T	G G T	
LEU CTT(	GIN GIN GLU CAACAAGAA( 1490	ARG THR GLY GTACAGGT 1550	GLU VAL 3 A A G T G 1610	TYR TAC	THR ACA
HIS GIN LEU	GIN G	ARG 1	CLU V	сых	GLY 1
PACCAACT	CAAC	C G T A	GAAG	ССТТ	G G A A
1430	1490	1550	1610	1670	1730
HIS	ARG	ASP	ASP	ILE	LEU
C A C	C G C C	G A T C	GAT(		TTG
VAL G T T (	MET ATG	ARG LEU : G C T T A ( 1540	ASN A A T	PHE GLY ILE GLY TYR GLY TTGGTATTGGTTACGGT 1660 1670	ASN PHE LEU GLY THR GLY ATTTCTTGGGAACAGGG 1720
LEU THR LTAACT (	GLU GAAA 1480	ARG C G C 1540	SER A G C 1 1600	PHE T T T 1660	ASN A A T 1720
LEU	GIN	ILE	GLY	ASN	ASP
TTA	C A G	A T T	GGT	A A C	G A T
ARG	ARG	LYS	ASN	ILE	GIN
CGTT	CGT(	A A A A	A A T	A T C P	C A A
NSP ALA GLY ARG ARG LEU THR VAL HIS GLN LEU ARG PHE GLA A T G C T G G C G C T T T A A C T G T T C A C C T T C G C T T T G A 1410 1420 1430	LEU FTTA 1470	VAL CLU LEU CLY LYS ILE ARG LEU ASP ARG THR CLY PHE PHE TTGAGTTAGGAAAATTCGCTTAGATCGTACAGGTTTCTT 1530 1560	ILE ' A T C 1590	SER 1 A G T A	NIA SER VAL LYS GIN ASP ASN PHE LEU GLY THR GLY ALA ALA CAAGTGTCAAACAAGATAATTTCTTGGGAACAGGGGGGGC 1710 1720 1730
G G A	THR ACT	LEU TTA	PRO C C T	G G T	WAL G T C
ALA	SER	alu	ASP	THR	SER
G C T	A G T	Ga G	G A T	A C G	AGTG
ASP GAT(	ASP GAT 50		G ILE : A A T T 1580	G ASN TAAC!	
VAL VAL A	ALA A	GIN LEU	ARG	ARG	GLN
STIGIIG	G C T G	CAATTA	C G A	C G T	C A A
1400	1460	15	15	16	17
VAL G T T	SER TCTG	GIN C A A	ASN A A C C	VAL LYS GLU STCAAAGAA( 1630	SER TYR GLN GTTATCAA 170
PHE	VAL	ASN SER	GLU	LYS	SER
TTT	GTT	A A T T C A (	GAA	A A A	AGT
ILE THR PHE	THR	ASN	THR VAL GLU	VAL	ILE
TAACCTTT	A C C C	A A T	ACAGTTGAAI	G T C	A T T
1390	1450	1510	1570	1630	1690
ILE	ASN	TYR	THR	LYS	SER GLY ILE
A T A	A A T	TAT	ACA	A A A	GTGGTATT?
ALA ILE THR PHE VAL VAL A AGCGATAACCTTTGTTGTTG 1390 1400	GLY ASN THR VAL SER ALA ASP SER THR LEU ARG GLN GLU MET ARG GLN GLU A G G A A A T A C C C T G T A C T T T A C G T C A G G A A A T G C G C C A A A G A A A G A A A G A A A G A A A B A B A A B	TRP TYR ASN SER GLN LEU V TTGGTATAATTCACAATTAG 1510 1520	GLU THR VAL GLU ASN ARG ILE ASP PRO ILE ASN GLY SER ASN ASP GLU VAL ASP VAL VAL CGAAACAGTTGAAATTGATCCTATCAATGGTAGCAATGAAGTGGATGTCGT 1570 1580 1590 1600	TYR LYS VAL LYS GLU ARG ASN THR GLY SER ILE ASN PHE GLY ILE CLY TYR GLY THR GLA A T A T A A A G A A C G T A A C A C G G G T A G T A T C A A C T T T G G T A C G G T A C A G A A G A A C G T A C G G G T A G T A C A C A C A A G A A C G T A C G G G T A G T A C A C A C T T T G G T A C G G T A C A G A A G A A C G T A C G G G T A G T A C A C T T T G G T T A C G G T A C A G A A G A A C G T A C G G G T A G T A C A C T T T G G T T A C G G T A C A C A C A C A C A C A C A C A C A	SER GLY ILE SER TYR GLN 1 GAGTGGTATTAGTTATCAAG 1690 1700

NS A

TRP ASN TYR

GEX

黑

SER

出

PHE ASP

LYS THR ASN ASP

ILE

GĽ

ASN

GLY

1990

2020

LYS

MET

GIN SER

TYR ILE

FE

ASI S

ARG

ASI

TYR

a B B

园

ALA

呂

S.

SER

当

TAAAATTAGTAACTTTGCTCTAGAATATAACCGTAATTTATATATTCAATGAAATT

TTAGGTTTCCCTGTAAATGAAATAACTCCTATTATGTAGGATTAGGCCATACTATAA

1950

1940

1930

1960

## FIG.1D.(CONTINUED)

		31/82	
GLU ; G A 1800	ASP : G A 1860	_	ASIN
THR	TYR	VAL	TYR
ACC	TAC	GTT	
TYR	ASN	ASN	景
TAT	A A C	A A T	
GCT O	GU GAA O	SER AGT 10	HIS
VAL SER ILE ALA CLY THR LYS ASN ASP TYR CLY THR SER VAL ASN LEU CLY TYR THR CLU	ASP GLY VAL SER LEU GLY GLY ASN VAL PHE PHE GLU ASN TYR ASP	THR SER SER ASN TYR LYS ARG THR THR TYR GLY SER ASN VAL THR	GLU ASN ASN SER TYR TYR VAL GLY LEU GLY HIS THR TYR
TAAGTATAGCTGGTACGAAAATGATTATGGTACGAGTGTCAATTTGGGTATCGA	3 A T G G T G T G G A A T G T T T T T G A A A C T A C G A	ACATCCTCTAAACTATAAGCGTACGACTTATGGAAGTAATGTTAC	
1750 1760 1770 1770 1780	1820 1830 1840 1850	1880 1990 1900 1900	
ASN	PHE	TYR	E
A A T	T T C	T A T	
VAL	VAL	THR	CILY
G T C	GTT	ACT	
SER	ASN	THR	VAL
A G T	A A T	A C G 2	
1780	1840	1900	
THR A C G	G G A	ARG C G T	TYR
G G T	GLY GGT	LYS A A G	TYR
TYR	LEU	TYR	SER
' T A T	FCTT	2 T A T	
1770	1830	1890	
ASP	SER	ASN	ASN
G A T	AGT	A A C	
ASN A A T	VAL G T A	SER	ASN
LYS	GLY	R SER	a a
A A A	G G T	:АТСС	
10	20	1880	
THR IACGA	ASP G G A T G 1820	THR ACA 181	ASN
G G T	TYR PHE THR LYS FATTTACTAAA( 1810	ASN SER LYS SER ASP ACTCTAAAAGTGATA 1870	PRO VAL
ALA	THR	SER	
G C T	ACT	AGT	
ILE	PHE	LYS	黑
A T A	T T T	A A A	
1750	1810	1870	
SER	TYR	SER	ריפת פרא
A G T	T A T	T C T	
VAL SER ILE ALA CLY THR LYS ASN ASP TYR CLY THR SER VAL ASN LEU CLY TYR THR CLY A G T A A G T A T A C C A A A A A A A G A T T A T G G T T A T A	PRO TYR PHE THR LYS ASP GLY VAL SER LEU GLY GLY ASN VAL PHE PHE GLU ASN TYR ASI GCCCTATTTTACTAAAGATGGTGTAAGTCTTGGTAATGTTTTCTTTGAAACTACGA 1810 1820 1830 1840	ASN SER LYS SER ASP THR SER SER ASN TYR LYS ARG THR THR TYR GLY SER ASN VAL THR TAACTCTAAAAGTGATACATCCTCTAACTATAAGCGTACGACTTATGGAAGTAATGTTAC 1870 1880 1890 1920	E

SUBSTITUTE SHEET

		32/82			
THR 2 A C 2160	PRO 2220		LEU 1 T T 2340	LYS ' A A 2400	GLU A G A 2460
VAL G T T <i>P</i> 21	TYR I'A (	ASN AATO	SER TCAT 23	ASN A A T	ALA 3 C A
ARG C G A (	PHE TTC'	ALA G C A	G G T	ASN A A T	SER A G C
GLY 3 G A (	GIN GLY 2 A G G G T 2	GLY TYR 3 G A T A T C 2270	ILE A T T 0	GLY GLN 3 G T C A A 1 2390	ALA G C T
GLY GLY 3 G T G G A 2150	GLN G C A G G 2210	GLY T G G A T 2270	THR ALA GLY GLY ILE GLY ACAGCGGTGGCATTGGT 2320 2330	сту с сстс 2390	THR , A C A G 2450
LEU CTT(	VAL 3 T A C	ALA 3 C A (	G G T	TYR GIN	ALA G C T
SER A G T (	ASP 3 A T (	SER	ALA 3 C G	TYR T A T	ILE A T C
ALA ; C A 1 2140	TYR LYS LEU SER ALA ASP VAL GIN GLY PHE TACAAACTAAGTGCAGATGTACAGGGTTTC 2190 2200 2210	ALA 3 C A T 2260	PHE TYR GIN THR TYR THR ALA GLY GLY ILE GLY SER LES TTCTATCAAACTTATACAGCGGGTGGCATTGGTTCATT 2310 2320 2330	ILE A T T ' 2380	ASN ALA ILE ALA THR ALA SER A A T G C A A T C G C T A C A G C T A G C C C C A S C C C C C A S C C C C C C C
LYS A A A G	SER A	LYS A A A O	TYR TAT	ALA G C A J	ASN A A T
GLY VAL 3 G G G T T 7 130	LEU	ALA G C A	THR TYR	ASN A A C G	CLY G G T
GLY 1 G G G 2130	TYR LYS LEU 'ACAAACTA 2190	SER \ T C T G 2250	TYR GLN 1 A T C A A 2310	PRO 3 C C T 2370	ILE GLY GLY ATTGGTGGT 1 2430
LYS A A A G 21	TYR TAC	VAL G T A '	TYR TAT	GLY GGGG 2	ILE ATT
THRACT	TYR 1 T A C	VAL G T T	PHE 3 T T C	ILE A T T	SER ASP VAL FCTGATGTG 2420
	LYS A A A 30	TRP T G G 10	$\sim$	SER A G C	ASP GAT
РНЕ 1 ТТСС 2120	ASN LYS A A C A A 1 2180	LEU TRP CTCTG(	LEU PRO TTACC 2300	αх s G G T A 2360	SER P TCTG 2420
ARG GLY TYR AGAGGCTAT 2110	ASP GAT1	ASP HIS	LYS ARG	ALA TYR GLY SER SCTTATGGTAGO 2360	SER AGT
C G C	SER	ASP G A T	LYS A A G	ALA G C T	ILE A T A
ARG A G A 2110	ДЛ G G Т 1 2170	ARG A G A G 2230	ASN A A C P 2290	PHE T T T ( 2350	ASN LYS ILE SER AATAAGATAAGT 2410
ASN A A T A	PRO C C A G	ASP GACA	PHE CLY	GGT	ASN A A T
LEU ASN ARG GLY TYR FHE FR CCTTAATAGAGGCTATTTCCC 2110 2120	ILE PRO GLY SER ASP ASN LYS TYR A A T T C C A G G T T C T G A T A C A A A T A C	LEU ASP ARG ASP HIS LÆU TRP VAL VAL SER ALA LYS ALA SER ALA GLY TYR ALA ASN GLY A T T A G A C A G A T C A C C T C T G G G T T G T A T C T G C A A A A G C A T C T G C A A A T G G A A A T G G A A A T G G A A A T G G A T A T	PHE CLY ASN LYS ARG LEU PRO TTTTGGAAACAAGCGTTTACC 2290 2300	ARG GLY PHE ALA TYR GLY SER ILE GLY PRO ASN ALA ILE TYR GLN GLY GLN ASN ASN LY. A C G C G T T T T G C T T A G C A T T G G G C C T A A C G C A A T T T A T C A A G G T C A A A T A A T A A A A 2350 2350 2360 2370 2400	PHE ASN LYS ILE SER SER ASP VAL ILE GLY GLY ASN ALA ILE ALA THR ALA SER ATTTAATAAGATAAGTTCTGATGTGATTGGTGGTAATGCAATCGCTACAGCTAGC 2430 2440 2450
υ	<b>A</b>	<b>«</b>	E→	Ø	K

		33/82			
LEU ; C T 2520	GLU 1 G A 2580	THR ' A C 2640	PRO 1 C C 2700	PHE 7 T T 2760	T T A 2820
SER TCCC 25	LEU	SER TCT1	LYS A A A	SER T C T	ATT
THR ACC	G G A	ALA G C C	ALA G C T	G G T	CTA
ARG C.G.A.A	ASN A A T	ARG C G C C	TYR TATO	GLY G G G O	T G C
VAL 1 3 T C C 2510	LYS ASN AAAAT 2570	THR 1 ACTC 2630	SER 1 TCTT 2690	ILE (ATTG 2750	стст 2810
THR A C A G	ASP G A T	ARG C G T	PHE T T T	SER A G T	G T T
E VAL SER ASP LYS SER GLN ASN THR VAL ARG THR SER LES TGTGAGTGATAAGAGTCAAATACAGTCCGAACCTCCT 2490 2520	SER T C A	1 PRO ASP TYR GLY LYS SER SER ARG THR ARG ALA SER THR ACCCGATTATGGCAAATCAAGCCGTACTCGCGCCTCTAC 2610 2620 2630 2640	PRO SER GLY PRO VAL VAL PHE SER TYR ALA CCTAGTGGACCAGTGGTATTTCTTATGCT 2670 2680 2690	PHE GLN PHE SER ILE GLY GLY TTCCAATTTAGTATTGGGGGT 2740 2750	GTCATCAGAACTCAAAACAACGTTCTCTGCCTAAT 2790 2800 2810
GIN C A A A 2500	LYS TRP LYS SER AAATGGAAATCA 2560	SER T C A A 2620	VAL G T G 2680	GLN C A A 2740	A A C 2800
SER AGTO	TRP T G G	LYS A A A S	PRO C C A (	雅 T T C C	AAA
LYS A A G	LYS A A A	TYR GLY FATGGC1 510	GLY GGA	GIN	ပ် မ
ASP C A T A 2490		TYR [ T A T 2610	SER 1 A G T 2670	VAL GLU G T C G A A ( 2730	1 G A A 2790
SER AGTO	ASN AAT	ASP G A T T 26	PRO C C T	O	T C A
VAL G T G	L TRP ASN THR TTGGAATACT 2550	PRO C C C	I SER ATCT	ASP TGAT	T C A
PHE TTT 0	VAL G T T		_	<u> </u>	
PRO PHE C C A T T 1 2480	SER VAL A G T G T 7 2540	ASP LEU G A C T T 1 2600	TRP GIN T G G C A 1 2660	ASN ASP A A T G A 1 2720	TTTT 2780
ACTO	ALA G C A J	LYS A A A G	GIN CAA'	ILE LYS LYS TYR GLU TTAAAAAATATGAA 2710	CTT
PRO C C A A	ALA G C G G	LEU TTG	PAE TTC(	TYR T A T	G A A
VAL G T G C 2470	ASP G A T G 2530	ASN VAL LEU ATGTCTTGA 2590	VAL GLY 3 T C G G A 7 2650	LYS A A A 2710	A T T 2770
ILE ATT G	VAL G T T	ASN A A T	VAL G T C	LYS A A A	*** T A A
LEU ILE VAL PRO THR PRO PHI GTTAATTGTGCCAACTCCATT 2470 2480	PHE VAL ASP ALA ALA SER VAL A T T T G T T G A T G C G G C A A G T G T 2530 2540	SER ASN VAL LEU LYS ASP LEG GAGCAATGTCTTGAAAGACTT 2590 2600	GLY VAL GLY PHE GIN TRP GLAAGGCAAGGTGTCGGATTCCAATGGCA	ILE LYS LYS TYR GLU ASN AS AATTAAAATATGAAAATGA 2710 2720	*** *** CTAATAAATTGAACTTTTTC 2780

	_	_
2		1
Ī	I	1
=		)
=	7	7
Ē	_	=
Ξ	_	-
	=	7
>	_	,
1	_	,
	_	:
L	_	1
т.		_

		34/82
A A A A A C 2880	G C A G C T 2940	34102
TTTATCAAATG 2870	СА G G C Т А Т G C Т 2930	
TAAGGATA 2860	: A C T T G C T T 2920	
AAACCATCATTTAATTAAGGATATTTATCAAATGAAAAC 2850 2850 2860 2870 2880	тбстттабстасасттсабстатстбстесабст 2910 2920 2930	TAATGCAGGTTATA 2970
A A T A T T A A A A A 2840	C C G C A C T T G C 2900	Сттттаттаа 2960
АТТ G G C A G A G A A A T A T T A 2830 2840	ATCGCCAAAGTAACCGCACT 2890 2900	6 А А G А А А А А Т Т G С Т Т Т Т А Т 2950 2960

FIG. 1E

JB-1042-9-4 DNA, PAK D15 IS THE SEQUENCE BEING TRANSLATED

		35/82			•
G G G A A A T T	AAAGGTGC	A T T A G T G T	тта сттт	A T C T G T T A	G A T T T T T T T 360
60	120	180	240	300	
А А А А Т А Т Т А	TCTATTGCA	A T G G C A T T G	G G C G G T C A T	G T G C A A A G C	ттатттаат
50	110	170	230	290	350
TTAACCTTG	G G C C A A T T	ттаа 6 ттт	G T A T T A G A T	T C T G A G C G G	G T G T T T G C A
40	100	160	220	280	340
CTTTCCCTT	A A T T T A A G T	G T G T A T T T T	C C A T T A C C A	АААССТGТ <u>Т</u>	A G C T T A A C G
30	90	150	210	270	330
A A C A G G A C A A	G T C A T T A A A T	TATTGGATTG	G A A T T T A T T T	TGTTAAAGGA	АСТ G Т Т А Т Т А
20	80	140	200		320
G C A T T G A A A A A A A C A G 10 20	CTGGCGATTT(	CATCAGCAAA' 130	A G G G A T T A T 190	C A A T G G A A G C 1 250	T T G G C G C A G C /
AAAAGO	ACTTAC	TGGTGC	АААТТТ	тттаас	TCGAAT

36/82						
ص 3	420	T G	I A G C 540	c G 600	c I G A 660	I A A A 720
LEU	4	ASP GAT 48	C A	PHE A TTCG 600	SER TCG	GLN C A A
LEU	<b>:</b> •	VAL G T G O	GLY G G T C	ARG CGAT	LYS A A A T	LYS A A A
SER A G T	• ) <del>!</del> .	ARG C G T (	ARG ALA : G T G C T C 530	SER GLY AGTGGT 590	ALA G C T	LEU CTT
ILE ALA	410	ILE A T T 470	ARG C G T 530	SER A G T 590	VAL G T G ( 650	ALA G C A 710
	) 	ASP G A T.A	PRO VAL	VAL G T A A	VAL G T T G	GLU GAA
LEU	; ,	LYS A A A (	PRO C C T	LEU PHE FTATTC ( 580	SER AGC(	THRACT
LEU CTT(	400	ALA G C A 1 460	LEU TTAC 520	LEU T T A 580	VAL VAL 3 T T G T T 7	PRO C C C 700
LYS A A A C		VAL G T G	SER A G T	ARG SER LEU PHE VAL SER GLY ARG PHE 1 CGCTCTTTATTCGTAAGTGGTCGATTCG 580 590	VAL G T T	ILE A T T
MET LYS LEU ATACAATCGATGAAAAACTT		PHE ALA ALA PRO PHE VAL ALA LYS ASP ILE ARG VAL ASP C TTTGCCGCACCTTTTGTGGCAAAGATATTCGTGTGGATG 450 460 470	GIN GIN ILE ARG ALA SER LEU PRO VAL ARG ALA GLY CAACAAATCCGAGCAAGTTTACCTGTTCGTGGT 510 520 530	ILE VAL ARG LTTGTCCGC7 570	GLU GLY ASP VAL LEU VAL VAL SER VAL VAL ALA LYS SER GAAGGCGATGTGCTTGTTGTTAGCGTTGTGGCTAAATCGA 630 660	LYS GLY ASN SER VAL ILE PRO THR GLU ALA LEU LYS GLN A A A A G G T A A C T C T G T T T T C C C A C T G A A G C A C T T A A A C A A A A C T T A A C A A A A
MET A T G	390	ALA PRO CACT 7	2 ARG C C G A 510	1 VAL TGTC 570	ASP VAL SATGTG (	г <u>стст</u> с стстс 690
T C G	χ. Υ	ALA G C A 4º	ILE ATC 5j	ILE ATT 57	ASP G A T	ASN AACT 690
CAA		ALA G C C	GIN C A A	ASN ILE VAL AATATTGTC 570	0 0 0 0	aly G G T
ATA				ALA G C T		LYS A A A
A G G	380	THR VAL CTGTG 440	LEU GLU TAGAA 500	ASP VAL SATGTG 560	HIS GIN SATCAA 620	ILE A T C 680
TAT	ŕ	THR ACT	LEU TTA	ASP G A T	HIS CAT	LYS A A A A
ΤΤΑ		THR THR THR ACGACAACGA	ASP G A C ?	ASP ASN 3 A C A A T (	VAL LYS ALA 3 T G A A A G C G (	ASP VAL ; A T G T T 1 ;70
AAT	370	THR A C A 430	GIN GLY AAGGT 490	ASP G A C 550	LYS A A A 610	ASP G A T 670
TAT		THR A C G	GLN C A A	THR ACTG	VAL G T G	SER TCAG
A C G T C T A T A A T T T A T A G G		HE GLY THR THR THR VAL TCGGTACGACACGACTGTG 430 440	LY VAL GIN GLY ASP LEU GLU GTGTTCAAGGTGACTTAGAA 490 500	RG VAL THR ASP ASN ASP VAL GTGTGACTGACAATGATGTG 550 550	SP ASP VAL LYS ALA HIS GIN A T G A T G T G A A G C G C A T C A A 610 620	LE ILE SER ASP VAL LYS ILE TCATTCAGATGTTAAAATC 670 680
A C		日 C	LY G T	RG G T C	SP A T C	LE T C

		3	7/82			
J P A T 780	С С 40	л д Т G 900	t L A T 960	G G 20	. T T 80	L A 'TA 1140
ASN GLU A A T G A A 71	GLU PGAAC BAAC 840	ASIN GLU ASP A NATGAAGATG 900	THR ACA 9	LYS PHE AAATTT 102	GLY G G C 10	LYS VAL AAGTT 114
ASN A A T	VAL GTTG	GLU GAA	SER A G T	LYS A A A	ASN A A T	LYS A A A
LEU T T A	ALA THR SCAACCO 830	ASN A A T	SER A G C	ASN A A 'F	ASN A A T	THR A C A
GLU LYS	ALA G C A 830	GIN ILE SAAATCI 890	SER A G T 950	GLY G G A 1010	LEU T T A 1070	LYS A A A 1130
GLU 3 A A S	ASN A A C G	GIN C A A	VAL G T T	TRP T G G	TYR T A T	G A A
ARG C G A G	TYR FAT	ILE A T T C	SER ICT	LEU	TYR I A T	ASP 3 A T
11.E A T T ( 760	ARG 2 G C 7 320	LEU TTA ? 880	GLU G A A 9	LYS A A A '	ASP 3 A T ' 360	ASN A A T (
LEU FTAA 7	GLY 3 G T (	ILE ATT	ASN A A C (	TRP r g G J	ARG C G T (	LEU CTA.
VAL 3 T T 1	ALA SER VAL GLY ARG TYR CAAGTGTAGGTCGCTATI 810 820	ALA GLU SCTGAAA	PHE LYS GLY ASN GLU SER VAL SER SER THR L TCAAGGGAACGAATCTGTTAGTAGCAGTACAT 930 940 950	ASP SER TRP TRP LYS LEU TRP GLY SATTCTTGGTGGAAATTATGGGGAA 990 1010	ILE A T T (	ASP VAL GIN LEU ASN ASP GLU LYS THR SATGTTCAGCTAAATGATGAAAAACAA 1110 1120
ASP G A T G 0	SER AGTO	ALA G C T (	LYS A A G (	SER TCTO	ALA G C A	WAL GTTO
GLY 3 G C ( 75	ALA G C A 2	ARG C G T G 870	FHE ITCA 930	ASP G A T 99	GLN C A G 105	ASP V G A T G 1110
IXS VAL GLY ASP VAL LEU ILE ARG GLU LYS LEU ASN GLU FA A A A A G T G G G A G A A A A A T G A A A T T C G A G A A A A T T A A A T T C G A G A A A A T T A A A T T 780 780	HIS TYR ALA SER VAL GLY ARG TYR ASN ALA THR VAL GLU F CACTATGCAAGTGTAGGTCGCTATAACGCAACCGTTGAAC 810 820 830	ASN ASN ARG ALA GLU ILE LEU ILE GIN ILE ASN GLU ASP P A A T A A T C G G G C T G A A A T T T T A A T T C A A A T C A A T G A G A T G 870 880 890	LEU THR PHE LYS GLY ASN GLU SER VAL SER SER THR L TTAACTTTCAAGGGAACGAATCTGTTAGTAGCAGTACAT 930 940 950	GIN PRO ASP SER TRP TRP LYS LEU TRP GLY ASN LYS PHE G CAACCTGATTCTTGGTGGAAATTATGGGGAAATATG 990 1020	ASP LEU GIN ALA ILE ARG ASP TYR TYR LEU ASN ASN GLY T GATCTGCAGGCAATTCGTGATTATTTAATTAATGGCT 1050 1060 1080	LYS THR ASP VAL GIN LEU ASN ASP GLU LYS THR LYS VAL 1 A A A A C G G A T G T T C A G C T A A A T G A T G A A A A A C A A A G T T A 1110 1120 1140
LYS A A A	HIS C A C	ASN A A T	LEU TTA	GLN C A A	ASP G A T	LYS A A A A
		PRO C C A 860		LEU T T A 980	LYS A A A A 1040	<b>5. .</b>
G G G	LYS A A A	LEU CTG	ALA G C A 1	GW GAA	GW GAG	ILE A T C
ALA ASN SCTAACO	VAL G T A	ASN THR AATACGO 350	LEU TTG	MET A T G	PE TTC	GLN C A A
ALA G C T 730	SER A G T 790	ASN A A T 850	LYS A A A 910	GIN C A A 970	ALA GLN PHE GLU LYS SCGCAATTCGAGAA/ 1030 1040	ALA 1 G C A 1090
ASP 3 A T	HE ALA LYS SER VAL LYS GLU TTGCCAAAAGTGTAAAAGAG 790 800	WAL 3 T C	SP LYS ALA LYS LEU ALA SER ATAAAGCAAAATTGGCATCA 910 920	EU GIN GIU GIN MET GIU LEU TACAAGAACAAATGGAATTA 970 980		YR ALA LYS ALA GIN ILE THR A T G C C A A A G C A C A A A T C A C T 1090 1100
LEU TTAC	ALA G C C	ILE ATT G	LYS A A A	GLN C A A	LU CLY AAGGTO	ALA G C C
N A	표 고 고	RO C T	SP A T	EU T A	LU A A	YR A T

_	_	_
(		2
Ī	1	Ī
=	_	Ī
=	_	7
Ē	-	=
Ξ	_	-
7	=	7
	_	く
1	_	ر
L	ı	i
	_	_
į	•	•
2		)
Ĺ	1	_

38/82					
LE G	LEU GLU PRO LEU LEU SER ALA LEU HIS LEU ASN A	ALA LYS LEU GLY G	YS T	ALA CLY ARG ARG LEU THR VAL ARG CLN LEU ARG P	SER THR LEU ARG GIN GLU MET ARG GIN GIN GIU G
TAG	TTGAACCTTTACTTCAGCATTACATTAAATG	SCAAACTTGGGG	A A A	CTGGACGATTTAACTGTTCGCCAACTTCGCT	\GTACTTTACGTCAGGAAATGCGACAACAAGAAG
1200	1230 1240 1260	1310 1320	1380	1410 1420 1420 1430	1470 1480 1490 1500
ARG ILE ILE GCATTATA 12	EU A FAA	EU G ITG	ALA ASN LYS SCAAATAAA 13	EU A	IN G A A G
I A C	II.	G E	A A A	C E	ש <sub>כ</sub>
ARG	HIS	LYS	ALA	GIN	GIN
C G O	C A '	A A	G C	C A 3	C A
ALA	LEU	ALA	ASP	ARG	ARG
G C A C	T T A	G C A	G A T G	C G C	C G A
1190	1250	1310	1370	1430	1490
ARG SER	ALA	LYS	ASP	VAL	MET
GTAGTO	G C A	A A A	G A C	G T T	A T G
ARG	SER	ILE	PRETTT	THR	GU
C G T	T C A	A T T		ACT	GAA
GLU GLY LEU GLN TYR ASP LEU SAAGGTTTACAGTATGACCTTC 1170	LEU A C T T 1240	ALA ASP VAL GLU ASN ALA ILE LYS 3 C A G A T G T A G A A A T G C A A T T A A A G A A 1300	PRO ASP PHE ASP CTGATTTTGACO	LEU T.T.A 420	GLN F C A G 1480
ASP	LEU	ASN	PRO	ARG	ARG
G A C	TTA	AAT	CCT	C G T	CGT
TYR T A T	PRO C C T	GLU GAA	ASN SER VAL AATTCTGTAC 1350	ARG C G A	LEU TTA
GIN	GU	VAL	SER	GLY	THR
C A G	GAA	G T A	T C T	G G A	ACT
0	O	0	0	0	0
LEU GI TTAC A	LEU (CTTG	ASP V G A T G 1290	ASN A A T 135	ALA G C T 141	SER 1 AGTA 1470
GLY	GLU	ALA	VAL	ASP	ASP
GGT	GAG	G C A	G T A	GAT(	GAT1
GW G	ALA	IIE	THR	VAL	ALA
	G C C G	A T T G	ACAC	GTT(	G C T G
ASN A A T 0 1160	SER T C T 1220	ASP G A T 7	THR A C A 1340	VAL G T T 1400	SER T C T ( 1460
ASP VAL	MET	SER	ASN	PHE	VAL
	A T G	A G T	A A C	TTT	G T T
ASP	GLY	ARG	GLY	THR	THR
G A T	G G T	C G T	GGT	A C C	A C C
R ILE CATT 1150	GLY 3 G G A 1210	ARG C G C 1270	TYR T A C	ILE A T A	ASN A A T 1450
SN VAL THR IIE ASP VAL ASN GLU GLY LEU GIN TYR ASP LEU ARG SER ALA ARG ILE ILE ( A T G T A A C C A T T G A T G A A G G T T T A C A G T A T G A C C T T C G T A G T G C A T T A T A G T G T A T A T A T A T A T	LY ASN LEU GLY GLY MET SER ALA GLU LEU GLU PRO LEU LEU SER ALA LEU HIS LEU ASN 1 GTAATCTGGGAGGTATGTCTGCCGAGCTTGAACCTTTACTTCAGCATTACATTAAATG 1210 1220 1230 1230	SP THR PHE ARG ARG SER ASP ILE ALA ASP VAL GLU ASN ALA ILE LYS ALA LYS LEU GLY G A T A C T T T C C G C C G T A G T A T T G C A G A T G T A G A A A T G C A A T T A A A G C A A A C T T G G G G 1270 1270 1280 1280 1280	LU ARG GLY TYR GLY ASN THR THR VAL ASN SER VAL PRO ASP PHE ASP ALA ASN LYS A A C G A G G T T A C A C A C A G T A A T T C T G T A C T G A T T T G A C G A T G C A A A T A A A A 1380 1380 1330 1340 1350 1350	HR LEU ALA ILE THR PHE VAL VAL ASP ALA CLY ARG ARG LEU THR VAL ARG CIN LEU CATTAGCGATAACCTTTGTTGATGCTGGACGACGTTTAACTGTTCGCCAACTT 1390 1430	HE GLU GLY ASN THR VAL SER ALA ASP SER THR LEU ARG GLN GLU MET ARG GLN GLU GLU (T.T.C.T.G.C.T.G.A.C.T.T.T.A.C.T.T.T.A.C.G.A.C.A.A.A.G.A.G.A.G.A.G.A.G.A.G.A
VAL	ASN	THR	ARG	LEU	GLU
G T A	A A T	ACT	C G A	T T A	G A A
SN A	LY G T A	SP A T	LU A A	C A	표단
•					

			/82		
LY P 3 T T 1560	ASP V 3 A T G 1620	LY T G T A 1680	A 3 G 740	/R T A T A 1800	ASN T AACT 1860
GLY G G 7	ASP G A 7	GEY G G J	GLY G G (	TYR T A :	ASN A A (
ARG THR GLY P : G T A C A G G T T 1560	ARG IILE ASP PRO IILE ASN GLY SER ASN ASP GLU VAL ASP V CGAATTGATCCTATCAATGGTAGCAATGATGAAGTGGATG	ARG ASN THR GLY SER ILE ASN PHE GLY ILE GLY TYR GLY T CGTAACACGGGTAGTATCAACTTTGGTATTGGTTACGGTA 1650 1660 1680	GIN THR SER ILE LYS GIN ASP ASN PHE LEU GLY THR GLY A CAAACAAGTATTAAACAAGATAATTTCTTGGGAACAGGGG 1710 1720 1730 1740	ASP TYR CLY THR SER VAL ASN LEU GLY TYR T SATTATGGTACGAGTGTCAATTTGGGTTATA 1780 1780 1790	ASP GLY VAL SER LEU GLY GLY ASN ILE PHE PHE GLU ASN 7 GATGGTGTAAGTCTTGGTGGAAATATTTTCTTTGAAAACT 1830 1840 1850
ARG C G T	GW GAA	G G T	G G A	leu TTG	PHETT
ARG LEU ASP GCTTAGATC 1550	ASP G A T 1610	ILE (ATTG) 1670	LEU T T G ( 1730	ASN A A T 1790	PHE T T C T 1850
LEU TTA	SER ASN	PHE CLY	ASN PHE	WAL G T C	ILE A T T
ARG C G C	SER A G C	PHETTT	ASN A A T	SER A G T	ASN A A T
ILE A A T T 1540	GLY r G G T 1600	ASN 2 A A C 1660	ASP G A T 1720	THR A C G 1780	GLY ' G G A 1840
VAL GLU LEU GLY LYS ILE FTGAGTTAGGAAAATTC 1530 1540	ILE ASN GLY ATCAATGGTA 1600	SER ILE ASN GTATCAACT 1660	LYS GIN ASP AACAAGATA 1720	GLY GGT	ASP GLY VAL SER LEU GLY GLY ASN ILE RATGGTAAGTCTTGGTGGAAATATT1 1830 1840
G G A	ILE A T C	SER A G T	LYS A A A	TYR T A T	LEUCTT
JU LEU AGTTA 1530	ILE ASP PRO TTGATCCTA 1590	R GLY : G G G T 1 1650	SER ILE AGTATTI 1710	ASP GAT 70	L SER 'AAGT 1830
GLU GAG 15	ASP E G A T C 1590	THR A C G G 1650	SER A G T	ASN A A T 17	VAL G T A
VAL GTT	ILE ATT	ASN A A C	THR	THR LYS ASN CGAAAAATG 1770	GGT
LEU	ARG C G A A	ARG C G T I	GAA	THR	ASP G A T
GIN C A A 1520	ASN A A C 1580	GLU G A A 1640	TYR T A T 1700	GLY ' G G T 1760	LYS ' A A A 1820
SER TCA	GW GAA	LYS	SER	ALA G C T	THR
ASN A A T	VAL GTT	WALGTO	ILE	ILE 'ATA'	PETT
TYR 3 T A T 1510	1 THR A A C A 1570	LYS F A A A 1630	GLY C G T 1690	SER A G T 1750	TYR C T A T 1810
LY THR TRP TYR ASN SER GIN LEU VAL GLU LEU GLY LYS ILE ARG LEU ASP ARG THR GLY I GAACTTGGTATAATTCACAATTAGTTGAGTTAGGAAAAATTCGCTTAGATCGTACAGGTT 1510 1520 1530 1530	HE PHE GLU THR VAL GLU ASN TCTTCGAAACAGTTGAAAACO 1570 1580	AL VAL TYR LYS VAL LYS GLU TCGTATATAAGTCAAAGAA(	HR GLU SER GLY ILE SER TYR CAGAGAGTGGTATCAGTTAT( 1690 1700	LA ALA VAL SER ILE ALA GLY THR LYS ASN ASP TYR CLY THR SER VAL ASN LEU GLY TYR 1 CGGCAGTAAGTATAGCTGGTACGAAAATGATTATGGTACGAGTGTCAATTTGGGTTATA 1750 1750 1760 1760 1770	HR GLU PRO TYR PHE THR LYS CCGAACCCTATTTTACTAAA( 1810 1820
THR ACT	PHE TTC	VAL G T A	GLU GAG	ALA G C A	G A A
LY G A A	品 T C	AL T C	C A	41 89 0	S C

_		4	0/82	
N V	R T	ET L	TYR A	>
TG	C T	F G A	A T A	
1920	1980	2040	2100	
AST.	TAC	MEI A T	TYF TA	ARC
SER	HIS	SER	ASN	GLY
A G T	C A T	TCA	A A C	
GLY	ប្ត ១	GLN	TRP	GLY
G G A	១ ១ ១	C A A	T G G	
TYR	LEU	ILE	GLY	ng i
T A T	T T A	A T T	G G T	
1910	1970	2030	2090	
THR ACT	GLY GGA	TYR T A T	PHETTT	SER
THR A C G	VAL G T A	LEU T.T.A	SER	ALA
ARG	R TYR	ASIN	PHE	LYS
C G T	T T A T	A A T	T T T	
900	1960	020	080	
THR SER SER ASN TYR LYS ARG THR THR TYR GLY SER ASN VACATCTCTAACTATAAGCGTACGACTTATGGAAGTAATG	ASN GLU ASN ASN SER TYR TYR VAL GLY LEU GLY HIS THR T AATGAAAATAACTCCTATTATGTAGGATTAGGCCATACCT 1950 1960 1970 1980	ALA LEU GLU TYR ASN ARG ASN LEU TYR ILE GIN SER MET L GCTCTAGAATATAACCGTAATTTATATATTCAATGA 2010 2020 2030	LYS THR ASN ASP PHE ASP PHE SER PHE GLY TRP ASN TYR A COAAACAAATGACTTTGATTTTCTTTTGGTTGGAACTATA N 2070 2070 2080 2100	PHE PRO THR LYS GLY VAL LYS ALA SER LEU GLY GLY ARG V
TYR	SER	ASN	PHE	(III)
T A T	T C C	A A C	TTT	
ASN A A C	ASN A A C	TYR TAT	ASP G A C 0	LYS
SER	ASN	GUU	ASN	景
TCT	AAT	GAA	A A T	
189	195	201	207	
SER	GU	LEU	THR	PRO
T C C	GAA	CTA	A C A	
THR	ASN	ALA	LYS	出
A C A	A A T	G C T	A A A	
F	VAL T A 1940			TYR
SER A G T	PRO C C T	ASN A A C	G G C	E S
LYS	PHE	SER	ASN	ARG
A A A	TTC	AGT	A A T	
ASN SER	GLY	ILE	GLY	ASN
ACTCT	\ G G T	A A T T	1 G G T	
1870	1930	1990	2050	
ASN	LEU	LYS	LYS	EE E
A A C	TTA	A A A	A A A	
YR ASP ASN SER LYS SER ASP	AL THR LEU GLY PHE PRO	YR ASN LYS ILE SER ASN PHE	YS PHE LYS GLY ASN GLY ILE	SN SER LEU ASN ARG GLY
ACGATAACTCTAAAAGTGAJ	TTACTTTAGGTTTCCCTG	ATAATAAATTAGTAACTTT	AATTTAAAGGTAATGGCATT	
1870 1880	1930	1990 2000	2050 2060	
YR	AL	YR	YS	<b>S S</b>
A C	T T	A T	A A	
		•		

TYR ALA A TTACTATTCCAGGTTCTGATAACAAATACTACAAACTAAGTGCAGATGTACAGGGTTCT A C C C A T T A G A C A G A T C A C C G C T G G G T T G T A T C T G C A A A A G C A T C T G C A G G A T A T G C A A VAL GIN GLY PHE ALA GLY 2210 SER ALA ASP SER ALA LYS ALA ASP ASN LYS TYR TYR LYS LEU HIS ARG TRP VAL VAL SER S. ASP ASP ARG ILE PRO GLY 13 混 A

A C A G C C T T A A T A G A G G C T A T T T C C C A A C T A A A G G G G T T A A A A G C C A A G T C T T G G T G G A C G A G

		41/8			
SN GLY PHE GLY ASN LYS ARG LEU PRO PHE TYR GIN THR TYR THR ALA GLY GLY ILE GLY S A T G G T T T T G G A A C A A C G T T C T A T C A A A C T T A T A C A G C G G T G G C A T T G G T T G T T G T T G T T G G T T G G T T G G T T G G T T G G T T G G T T G G T T T T G G T T T G G T T T G G T T T G G T T T T G G T T T T G G T T T T G G T T T T G G T T T T T G G T T T T G G T T T T G G T T T T T G G T T T T G G T T T T G G T T T T T G G T	PRO ASN ALA ILE TYR ALA GLU HIS GLY A CCTAATGCAATTTATGCCGAACATGGTA 2380 2380 2400	SN GLY THR PHE ASN LYS ILE SER SER ASP VAL ILE GLY GLY ASN ALA ILE THR THR ALA S A T G G T A C T T T T A A G A T A A G T T C T G A T T G G T G G T A A T G C A A T C A C A G C G A 2410 2420 2460	ER ALA GLU LEU ILE VAL PRO THR PRO PHE VAL SER ASP LYS SER GLN ASN THR VAL ARG T GTGCAGAACTTATTGTACCAACTCCATTTGTGAGTGATAAAAGCCAAAATACAGTCCGAA 2470 2520	HR SER LEU PHE VAL ASP ALA ALA SER VAL TRP ASN THR LYS TRP LYS SER ASP LYS ASN G CCTCCCTATTTGTTGATGCGGCAAGTGTTTGGAATACTAAATGGAAATCAGATAAAAG 2530 2530 2530	LY LEU GLU SER LYS VAL LEU LYS ASP LEU PRO ASP TYR GLY LYS SER SER ARG ILE ARG A GATTAGAGAGCAAGGTCTTGAAAGACTTACCTGATTATGGCAAATCAAGCCGTATTCGCG 2590 2630 2640
IIE	HIS	THR C T	VAL	LYS	ILE
\ T T (	CAT		3 T C	A A A i	A T T
GEN	GLU	THR	THR	ASP	ARG
1 G C 1	3 A A C	ACAI	A C A (	3 A T A	C G T
GLY GLY	R ALA	ILE THR THR	ASN THR VAL	SER	SER
5 G T G G C /	1 T G C C C	ATCACAACT(	AATACAGTC	I C A G	A G C C
2330	2390	2450	2510	2570	2630
ALA	TYR	ALA	GIN	LYS	SER
3 C G G	' A T G	; C A A	C A A A	A A A 7	CCA1
THR	ILE	ASN	SER	TRP	LYS
CAG	TTT	ATG	G C C	' G G /	A A A 7
TYR THR	ALA	ASP VAL ILE CLY CLY ASN ALA	LYS SER	SER VAL TRP ASN THR LYS TRP LYS	ASP LEU PRO ASP TYR GLY LYS SER
FATACA(	F G C A A	ATGTGATTGGTGGTAATGCAA	FAAAAGC(	GTGTTTGGAATACTAAATGGAAA?	RACTTACCTGATTATGGCAAATCA
320	2380	2430	2500	2550	2610 2610
THR TYR ACTTA 2320	ASIN NATG 23	CLY 3 G T G 24	ASP G A T A 25	THR ACTA	TYR F A T ( 26
GIN	PRO	ILE	SER	ASN	ASP
: A A A	CTP	A T T C	AGTO	A A T 1	3 A T ?
TYR GIN FATCAA!	ILE GLY PRO .TTGGGCCT1 2370	VAL S T G I	PHE VAL TTGTG1 2490	TRP FGG1	PRO C C T (
РНЕ Т	ILE GATO 2370	ASP	PHE V	VAL 1	LEU 1
ГТСТ		3 A T (	LTTG	3 T T T	TTAC
2310		243(	2490	2550	2610
PRO	SER ILE GLY AGTATTGGG 2370	SER I C T (	PRO C C A 7	SER A G T	ASP G A C
LEU	GLY	SER	THR	ALA	LYS
PTAC	G G T I	AGTT	ACT(	G C A A	A A A G
ARG	TYR	ILE	PRO	ALA	LEU
C G T 1	T A T (	A T A 1	C C A 2	G C G (	T T G
2300	2360	2420	2480	2540	2600
LYS	ALA	LYS	VAL	ASP	WAL
A A G (	3 C T '	A A G	G T A	G A T	G T C
GLY ASN 3 G A A A C 1 290	PAE T T T	ASN A A T	ILE A T T	VAL G T T	LYS A A G
GLY	ARG GLY PHE ALA TYR	PHE	ALA GLU LEU ILE VAL	PHE	SER
T G G A 2	3 G C G G T T T T G C T T A T	T T T	SCAGAACTTATTGTAC	\ T T T	5 A G C
2290	2350 2350	410	2470	2530	2590
РИЕ ГТТС 22	ARG CGC	THR ACT	GW GAA 2	LEU CTA	GLU GAG
SN GLY ATGGTT	ER LEU ARG GLY PHE ALA TYR GLY CATTACGCGTTTTGCTTATGGT 2350 2350	SN GLY THR PHE ASN LYS ILE ATGGTACTTTTAATAAGAT1 2420	ALA 3 C A	HR SER LEU PHE VAL ASP ALA CCTCCCTATTTGTTGATGCG 2530	LY LEU GLU SER LYS VAL LEU GATTAGAGAGCAAGGTCTTC 2590
SN A T (	C A	NS A	E C T C	版 い	LY G A

R A 1 T G 2700	9 9 9
TYR T A T 27(	GLY (G G G G G 2760
SER	ile
I C T	a t t
GLN TRP GLN SER PRO ILE GLY PRO LEU VAL PHE SER TYR ACANTGCANTTTCTTATG 2A CCANTGGTATTTCTTATG 2700 2670 2690 2700	R GLU ASN ASP ASP VAL GLU GIN PHE GLN PHE SER ILE GLY C TGAAAATGATGTCGAACAGTTCCAATTTAGTATTGGGG 20 2730 2740
VAL	PHE
G T A	T T T
2690	2750
LEU	GLN
TTG	C A A
PRO	PHE
C C A	TTC
GLY	GLN
G G A	1 C A G
2680	2740
ILE ATT(	VAL GLU GIN STCGAACAC 2740
PRO	WAL
C C T	G T C
SER TCT	ASP G A T 10
GIN	ASP
C A A T	G A T G
2670	2730
TRP	ASN
TGG	A A T
GIN	GLU
C A A	GAA
္မပ္	TYR T A T 2720
G G A	LYS A A A
WAL	LYS
G T C	A A A
GLY	11.E
G G T	, A T T
2650	2710
LA SER THR GLY VAL GLY PHE	LA LYS PRO ILE LYS LYS TYR
CCTCTACAGGTGTCGGATT	CTAAACCAATTAAAAATA
2650 266	2710 272
SER	LYS
T C T	A A A (
C C I	LA C T /

		42	/82
	TTCTCTGC	2820	·
	CAAAAACGACGTTCTCTGC	2810	
	TCGTCATCAGAACTC	2800	
	TTTTTCGTCA	2790	
	CTCTTTCTAATAAATTGAACT	2780	
**	ATAA		
* * *	TAI	2770	
器	r T T (		
SE	TCI		
Ľ	O S		

•	
TCAAAT	286
AAGGATATTTA	2870
CATTTAATT	2860
A T T A A A C C C A T (	2850
CAGAGAAAAT	2840
TAATTGAATTGGG	2830

GCTATGC	2940
TTGCTTCAG	
AGGTTTTGCAC	2920
ACTTGCTTTA	2910
GTAACCGC	2900
GAAAACATCGCAAAA	2890

TTTTTCAA	
GTTATA	2980
CATTAATGCAG	2970
AATTGCTTT	2960
GCTGAAGAAAA	2950
TTCC	

FIG. 1F

		1 ACAGGACAGCTTTCCCTTTTAACCTTGAAAATATTAGGC	1 aaaaGGCATTGAAAAAACAGGACAACTTTCCCTTTTAACCTTGAAATATTAGGC	1 GGCATTGAAAAACAGGACAGCTTTTCCCTTTTTAACCTTGAAAATATTAGGC	aaaaggcattgaaaaaacaggacagctttcccttttaaccttgaaaatattaggg	
<b>←</b> -1	<b>~</b>	₹~7				
cad15	minnad15	eagand15	pakd15	sb33d15	consensus	
	cad15 1	cad15 1 minnad15 1	1 1	1 1 ACAGGACAGCTT 1	1 1 1 5 1 aaaaGGCATTGAAAAA 1 GGCATTGAAAAA	1 1 1 aaaaGGCATTGAAAAA 1 GGCATTGAAAAA 3 aaaaggcattgaaaaa

43/82

CTTACTGGCGATTTGTCATTAAATTAATTTAAGTGGGCCAATTTCTATTGCAAAAGGTGCTG	CTTCCTGGCGATTTGTCATTAAATAATTTAAGTGGGCCAATTTCTATTGCAAAAGGTGCTG			cttactggcgatttgtcattaaataatttaagtgggccaatttctattgcaaaaggtgctg	44/82	GCaCATCAGCAAATATTGGATTGGTGTATTTTTAAGTTTTTATGGCACTGATTAGTGTAAA	07 GCcCATCAGCAAATATTGGATTGGTGTATTTTTTAAGTTTTTATGGCACTGATTAGTGTAAA			gcgcatcagcaaatattggattggtgtattttttaagtttttatggca-tgattagtgtaaa
9	46 (	62	28		₽	29	107	123	119	
minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus

					45	/82				
1 	128 TTTAGGGATTATGAATTTATTTCCATTACCAGTATTAGATGGCGGTCATTTAGTTTTTTAA		184 TTTAGGGATTATGAATTTATTTCCATTACCAGTATTAGATGGCGGTCATTTAGTTTTTT		tttagggattatgaatttatttccATTACcagtattagatggcggtcatttagttttttta		189 ACAATGGAAGCTGTTAAAGGAAAACCTGTTTCTGAGCGGGTGCAAAGCATCTGTTATCGA	229 ACAATGGAAGCTGTTAAAGGAAAACCTGTTTCTGAGCGGGTGCAAAGCATCTGTTATCGAA	245 ACAATGGAAGCTGTTAAAGGAAAACCTGTTTCTGAGCGGGTGCAAAGCATCTGTTATCGA	241 ACAATGGAAGCTGTTAAAGGAAAACCTGTTTCTGAGCGGGGGCAAAGCATCTGTTATCGAA
-	12	168	18	180		7	18	22	24	24
cad15	minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15

consensus

acaatggaagctgttaaaggaaaacctgtttctgagcgggtgcaaagcatctgttatcgaa

gccaagCTTAACGGTGTTTGCATTAATTAATGATTTTACGTCT	0 TTGGCGCACCACTGTTATTAAGCTTAACGGTGTTTGCATTATTTAATGATTTTTACGTCT		306 TTGGCGCACTGTTATTAAGCTTAACGGTGTTTGCATTATTAATGATTTTTACGTCT  100 TTGGCGCACTGTTATTAAGCTTAACGGTGTTTGCATTATTAATGATTTTTACGTCT  100 TTGGCGCACTGTTATTAAGCTTAAGCTTAACGGTGTTTAATGATTTTTAAGGTTCT  100 TTGGCGCACTGTTATTAAGCTTAAGCTTAAGGTTTATTAAGGTTTTTAAGGTTCT  100 TTGGCGCACTGTTATTAAGCTTAAGGTTTAAGGTTTAAGGTTTAAGGTTCTAAGAGATTAAGAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA	11111111111111111111111111111111111111	N ttggcgcagcactgttattAAGCTTAACGGTGTTTTGCATTATTTAATGATTTTTTACGTCT	52 ATAATTTATATAGGATACAATCGATGAAAAACTTCTAATCGCAAGTTTATTATTCGGTAC	1 ATAATTTATATAGGATACAATCGATGAAAAAACTTCTAATCGCAAGTTTATTATTCGGTAC	SI ATAATTTATATAGGATACAATCGATGAAAAAACTTCTAATCGCAAGTTTATTATTCGGTAC	367 ATAATTTATATAGGATACAATGGATGAAAAACTTCTAATCGCAAGTTTATTATTCGGTGC	
	250	290		302			311	351		
cad15	minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15	pakd15	

					47/8	2					
ATAATTTATATAGGATACAATGGATGAAAAACTTCTAATCGCAAGTTTATTATTCGGTAC	ATAATTTATATAGGATACAATGGATGAAAAAACTTCTAATCGCAAGTTTATTATTCGGTAC	113 GACAACGACTGTGTTTGCCGCACCTTTTGTGGCAAAAGATATTCGTGTGTGGATGTTCAA	GACAACGACTGTGTTTGCCGCACCTTTTGTGCCAAAGATATTCGTGTGTGGATGGTGTTCAA		GACAACGACTGTGTTTGCCGCACCTTTTGTGcCAAAAGATATTCGTGTGTGGTGTTTCAA	GACAACGACTGTGTTTGCCGCACCTTTTGTGGCAAAAGATATTCGTGTGGATGGTGTTCAA	GACAACGACTGTGTTTGCCGCACCTTTTGTGGCAAAAGATATTCGTGTGGATGGTGTTCAA		GGTGACTTAGAACAAATCCGAGCAAGTTTACCTGTTCGTGCCGGTCAGCGTGTGACTTG	473 GGTGACTTAGAACAAATCCGAGCAAGTTTACCTGTTCGTGCCGGTCAGCGTGTGTGT	
363		113	372	412	428	424		174	433	473	
sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15	

485 485 235 235 534 550 550 296 296	489 GGTGACTTAGAACAAATCCGAGCAAGTTTACCTGTTCGTGCtGGTCAGCGTGTGACTG	GGTGACTTAGAACAAATCCGAGCAAGTTTACCTGTTCGTGCcGGTCAGCGTGTGACTG	ACAATGATG	550 ACAATGATGTGGCTAATATTGTCCGCTCTTTATTCGTAAGTGGTCGATTCGATGTGATGTGAA  [	ACAATGATGTGGCTAATATTGTCCGCTCTTTATTCGTAAGTGGTCGATTCGATGTGAA	296 AGCGCATCAAGAAGGCGATGTGCTTGTTGTTAGCGTTGTGGCTAAATCGATCATTTCAGAT	555 AGCGCATCAAGAAGCCGATGTTTTTTTTTTAGCGTTGTGGCTAAATCGATCATTTCAGAT	
pakd15 sb33d15 consensus eagand15 pakd15 sb33d15 consensus cad15 minnad15		sensus	10	pakd15	ısensus	cad15		

595 AGCGCATCAAGAAGGCGATGTGCTTGTTAGCGTTGTGGCTAAATCGATCATTTCAGAT	611 AGCGCATCAAGAAGGCGATGTTGTTTGTTAGCGTTGTGGCTAAATCGATCATTTCAGAT	607 AGCGCATCAAGAAGGCGATGTGCTTGTTAGCGTTGTGGCTAAATCGATCATTTCAGAT	AGCGCATCAAGAAGGCGATGTGCTTGTTAGCGTTGTGGCTAAATCGATCATTTCAGAT	357 GTTAAAATCAAAGGTAACTCTGTTATTCCCACTGAAGCACTTAAACAAAACTTAGATGCTA	616 GTTAAAATCAAAGGTAACTCTGTTATTCCCACTGAAGCACTTAAACAAAC	656 GTTAAAATCAAAGGTAACTCTGTTATTCCCACTGAAGCACTTAAACAAAC	672 GTTAAAATCAAAGGTAACTCTGTTATTCCCACTGAAGCACTTAAACAAAC	668 GTTAAAATCAAAGGTAACTCTATTATTCCACCTGAAGCACTAAAACAAAACTTAGATGCTA	GTTAAAATCAAAGGTAACTCTgTTATTCCcaCTGAAGCACTtAAACAAAACTTAGATGCTA
	pakd15	sb33d15	snsı	cad15			pakd15	sb33d15	sns
eagand15	pa	sb3	consensus	υ	minnad15	eagand15	pa	sb3	consensns

49/82

cad15

ACGGGTTTAAAGTTGCCGATGTTTTAATTCGAGAAAATTAAATGAATTTGCCAAAAGTGT	ACGGGTTTAAAGTTGCCGATGTTTTAATTCGAGAAAATTAAATTGAATTTGCCAAAAGTGT	ACGGGTTTAAAGTTGGCGATGTTTTAATTCGAGAAAATTAAATGAATTTGCCAAAAGTGT		ACGGGTTTAAAGTTGGCGATGTTTTAATTCGAGAAAAATTAAATGAATTTGCCAAAAGTGT	∢-	AAAAGAGCACTATGCAAGTGTAGGTCGCTATAACGCAACAGTTGAACCTATTGTCAATACG	AAAAGAGCACTATGCAAGTGTAGGTCGCTATAACGCAACAGTTGAACCTATTGTCAATACG	AAAAGAGCACTATGCAAGTGTAGGTCGCTATAACGCAACCGTTGAACCTATTGTCAATACG		AAAAGAGCACTATGCAAGTGTAGGTCGCTATAACGCAACaGTTGAACCTATTGTCAATACG
677	717	733	729		479	738	778	794	790	
minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus

	cad15	540	CTACCAAATAATCGCGCTGAAATTTTAATTCAAATCAATGAAGATGATAAAGCAAAATTGG
	minnad15	799	CTACCAAATAATCGCGCTGAAATTTTAATTCAAATCAATGAAGATGATGATAAAGCAAAATTGG
	eagand15	839	CTACCAAATAATCGCGCTGAAATTTTAATTCAAATCAATGAAGATGATAAAGCAAAATTGG
	pakd15	855	CTGCCAAATAATCGLGCTGAAATTTTAATTCAAATCAATGAAGATGATAAAGCAAAATTGG
	sb33d15	851	
	consensus		U CTaCCAAATAATCGcGCTGAAATTTTAATTCAAATGAAGATGATAAAGCAAAATTGG ©
	cad15	601	
	minnad15	860	860 CATCATTAACTTTCAAGGGGAACGAATCTGTTAGTAGCAGTACATTACAAGAACAAATGGA
	eagand15	006	CATCATTAACTTTCAAGGGGAACGAATCTGTTAGTAGCAGTACATTACAAGAACAAATGGA
	pakd15	916	CATCATTAACTTTCAAGGGGAACGAATCTGTTAGTAGCAGTACATTACAAGAACAAATGA
	sb33d15	912	
•	consensus		CATCATTAACTTTCAAGGGGAACGAATCTGTTAGTAGCAGTACATTACAAGAACAAATGGA

662 ATTACAACCTGATTCTTGGTGGAAATTATGGGGAAATAAAT	921 ATTACAACCTGATTCTTGGTGGAAATTATGGGGAAATTTGAATTTGAAGGTGCGAATTCGAG	961 ATTACAACCTGATTCTTGGTGGAAATTATGGGGAAATTAAATTTGAAGGTGCGCAATTCGAG	977 ATTACAACCTGATTCTTGGTGGAAATTATGGGGAAATTAATAAATTTGAAGGTGCGCAATTCGAG	973 ATTACAACCTGATTCTTGGTGGAAATTATGGGGAAATAAAT	U ATTACAACCTGATTCTTGGTGGAAATTATGGGGAAATATTGAAGGTGCGCAATTCGAG N © N	723 AAAGATTTGCAGTCAATTCGTGATTATTTTAAATAATGGCTATGCCAAAGCACAAATTA	982 AAAGATTTGCAGTCAATTCGTGATTATTTTAAATAATGGCTATGCCAAAGCACAAATTA	1022 AAAGATTTGCAGTCAATTCGTGATTATTTAAATAATGGCTATGCCAAAGCACAAATTA	1038 AAAGATCTGCAGGCAATTCGTGATTATTTTAAATAATAATGGCTATGCCAAAGCACAAATCA	II ATtT	AAAGATETGCAGECAATTCGTGATTATTTTAAATAATGGCTATGCCAAAGCACAAATEA
cad15	minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15 1	pakd15 1	sb33d15 1	consensus

	CTAAAACGGATGTTCAGCTAAATGATGAAAAAACAAAAGTTAATGTAACCATTGATGTAAA	CTAAAACGGATGTTCAGCTAAATGATGAAAAAAAAAAAA	CTAAAACGGATGTTCAGCTAAATGAAAAAAAAAAGTTAATGTAACCATTGATGTAAA	CTAAAGCGGATGTTCAGCTAAATGATGAAAAAAACAAAAGTTAATGTAACCATTGATGTAAA O	CTAAAACGGATGTTCAGCTAAATGATGAAAAAAAAAAGTTAATGTAAACCATTGATGTAAA W SOOTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	TGAAGGTTTACAGTATGACCTTCGTAGTGCACGCATTATAGGTAATCTGGGAGGTATGTCT	TGAAGGTTTACAGTATGACCTTCGTAGTGCACGCATTATAGGTAATCTGGGAGGTATGTCTTTTTTATAGGTAATCTGGGAGGTATGTCTTTTTTTT	TGAAGGTTTACAGTATGACCTTCGTAGTGCACGCATTATAGGTAATCTGGGAGGTATGTCT	TGAAGGTTTACAGTATGACCTTCGTAGTGCACGCATTATAGGTAATCTGGGAGGTATGTCT		TGAAGGTTTACAGTATGACCTTCGTAGTGCACGCATTATAGGTAATCTGGGAGGTATGTCT
784	1043	1083	1099	1095		845	1104	1144	1160	1156	
	minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus

906 GCCGAGCTTGAACCTTTCAGCATTACATTTAAATGATACTTTCCGCCGTAGTGATA			1 GCCGAGCTTGAACCTTTCAGCATTACATTTAAATGATACTTTCCGCCGTAGTGATA 1 GCCGAGCTTGAACCTTTCAGCATTACATTTAAATGATACTTTCCGCCGTAGTGATA		U GCCGAGCTTGAACCTTTACTTTACATTTAAATGATACTTTCCGCCGTAGTGATA & © © © © © © © © © © © © © © © © © ©	967 TTGCAGATGTAGAAAATGCAATTAAAGCAAAACTTGGAGAACGCGGTTACGGTAGCGCAAC		TTGCAGATG1	TTGCAGATG1	TTGCAGATG	TTGCAGATGTAGAAAATGCAATTAAAAGCAAAACTTGGAGAACGCGGTTACGGTTACCAAAAC
90	1165	1205	1221	1217		96	1226	1266	1282	1278	
cad15	minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus

GATGCTGGACGACGTTTAACTGTTCACCTTTCGCTTTTGAAGGAAATACCGTTTCTGCTG GATGCTGGACGACGTTTAACTGTTCGCCAACTTCGCTTTGAAGGAAATACCGTTTCTGCTG	1400	consensus
GATGCTGGACGACGTTTAACTGTTCGCCAACTTCGCTTTGAAGGAAATACCGTTTCTGCTG	_	pakd15
GATGCTGGACGACTTTAACTGTTCGCCAACTTCGCTTTGAAGGAAATACCGTTTCTGCTG		eagand15
GATGCTGGACGACGTTTAACTGTTCGCCAACTTCGCTTTGAAGGAAATACCGTTTCTGCTG		
GATGCTGGACGTTTAACTGTTCGCCAACTTCGCTTTGAAGGAAATACCGTTTCTGCTG		_
5/82		
gGTAAATTCaGTACCTGATTTTGAtGATGCAAATAAAACATTAGCGATAACCTTGTTGTT		consensus
<b>AGTAAATTCTGTACCTGATTTTGACGATGCAAATAAAACATTAGCGATAACCTTTGTTGTT</b>	1339	sb33d1.5
	1343	pakato
		T 48 - 1
GGTAAATTCAGTACCTGATTTTGATGATGCAAATAAAACATTAGCGATAACCTTGTTGTT	1327	eagand15
GGTAAATTCAGTACCTGATTTTGATGATGCAAATAAAACATTAGCGATAACCCTTGTTGTT	1287	minnad15
cadis 1028 GGTAAATTCAGTACCTGATTTTGATGATGCAAATAAAACATTAGCGATAACCCTTGTTGTT	1028	cadl5

ATAGCACTTTACGTCAGGAAATGCGCCAACAAGAAGGAACTTGGTATAATTCACAATTAGT		ATAGTACTTTACGTCAGGAAATGCGACAACAAGAAGGAACTTGGTATAATTCACAATTAGT	ATAGTACTTTACGTCAGGAAATGCGCCAACAAGAAGGAACTTGGTATAATTCACAATTAGT	ATAGCACTTTACGTCAGGAAATGCGCCAACAAGAAGGAACTTGGTATAATTCACAATTAGT U O O	N TGAGTTAGGAAAAATTCGCTTAGATCGTACAGGTTTCTTCGAAACAGTCGAAAACCGAATT	TGAGTTAGGAAAATTCGCTTAGATCGTACAGGTTTCTTCGAAACAGTCGAAATT	TGAGTTAGGAAAATTCGCTTAGATCGTACAGGTTTCTTCGAAACAGTCGAAACCGAATT	TGAGTTAGGAAAATTCGCTTAGATCGTACAGGTTTCTTCGAAACAGTTGAAACCGAATT	TGAGTTAGGAAAAATTCGCTTAGATCGTACAGGTTTCTTCGAAACAGTTGAAAACCGAATT	TGAGTTAGGAAAAATTCGCTTAGATCGTACAGGTTTCTTCGAAACAGTcGAAAACCGAATT
1150	1449	1465	1461	·	1211	1470	1510	1526	1522	
cad15 minnad15 1	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus

GATCCTATCAATGGTAGTAATGATGAGTGGATGTCGTATATAAAGTCAAAGAACGTAACA	GATCCTATCAATGGTAGTGATGATGTCGTATATAAAGTCAAAGAACGTAACA	SATCCTATCAATGGTAGTGATGATGTGGATGTCGTATATAAAGTCAAAGAACGTAACA	GATCCTATCAATGGTAGCAATGATGATGTCGTATATAAAGTCAAAGAACGTAACA	GATCCTATCAATGGTAGCAATGATGAGTGGATGTCGTATATAAAGTCAAAGAACGTAACA	GATCCTATCAATGGTAGTAATGATGAGTGGATGTCGTATATAAAGTCAAAGAACGTAACA U	32	CGGGTAGTATCAACTTTGGTATTGGTTACGGTACAGAGTGGTATTAGTTATCAAGCAAG	CGGGTAGTATCAACTTTGGTATTGGTTACGGTACAGAGTGGTATTAGTTATCAAGCAAG	CGGGTAGTATCAACTTTGGTATTGGTTACGGTACAGAGTGGTATTAGTTATCAAGCAAG	CGGGTAGTATCAACTTTGGTATTGGTTACGGTACAGAGTGGTATCAGTTATCAAACAAG	CGGGTAGTATCAACTTTGGTATTGGTTACGGTACAGAGAGTGGTATTAGTTATCAAGCAAG	CGGGTAGTATCAACTTTGGTATTGGTTACGGTACAGAGAGTGGTATtAGTTATCAAgCAAG
1272	1531	1571	1587	1583			1333	1592	1632	1648	1644	
cad15	minnad15 1	eagand15 1	pakd15	sb33d15	consensus		cad15	minnad15	eagand15	pakd15 1648	sb33d15	consensus

GATTATGGTACGAGTGTCAATTTGGGTTATACCGAGCCCTATTTTACTAAAGATGGTGTAA

consensus

TGTTAAACAAGATAATTTCTTGGGAACAGGGGCGGCAGTAAGTA	TGTTAAACAAGATAATTTCTTGGAACAGGGCGGCAGTAAGTA	TGTTAAACAAGATAATTTCTTGGGAACAGGGGCGCCAGTAAGTA		TGTCAAACAAGATAATTTCTTGGGAACAGGGGCGGCAGTAAGTA	TGT LAAACAAGATAATTTCTTGGGAACAGGGGGGGGGGATAAGTATAGCTGGTACGAAAAAT U ©	SATTATGGTACGAGTGTCAATTTGGGTTATACCGAGCCCTATTTTACTAAAGATGGTGTAAA	GATTATGGTACGAGTGTCAATTTGGGTTATACCGAGCCCTATTTTACTAAAGATGGTGTAAA	GATTATGGTACGAGTGTCAATTTGGGTTATACCGAGCCCTATTTTACTAAAGATGGTGTAAA		GATTATGGTACGAGTGTCAATTTGGGTTATACCGAGCCCTATTTTACTAAAGATGGTGTAA
1394	1653	1693	1709	1705		1455	1714	1754	1770	1766
cad15 1394	minnad15	eagand15	pakd15 1709	sb33d15	consensus	cad15	minnad15	eagand15	pakd15 1770	sb33d15

1516 GTCTTGGTGGAAATGTTTTCTTTGAAAACTACGATAACTCTAAAAGTGATACATCCTCTAA	GTCTTGGTGGAAATGTTTTCTTTGAAAACTACGATAACTCTAAAAGTGATACATCCTCAA	GTCTTGGTGGAAATGTTTTCTTTGAAAACTACGATAACTCTAAAAGTGATACATCCTCTAAA	GTCTTGGTGGAAATATTTTTTTTTTTTTTTTTTTTTTTT	GTCTTGGTGGAAATGTTTTTTTTTTTTTTTTTTTTTTTT	GTCTTGGTGGAAATGTTTTCTTTGAAAACTACGATAACTCTAAAAGTGATACATCCTCTAA U	8 S S S S S S S S S S S S S S S S S S S		_G_	CTATAAGCGTACGACTTACGGAAGTAATGTTACTTTAGGTTTCCCTGTAAATGAAAATAAC	CTATAAGCGTACGACTTATGGAAGTAATGTTACTTTAGGTTTCCCTGTAAATGAAAATAAC	CTATAAGCGTACGACTTATGGAAGTAATGTTACTTTAGGTTTCCCTGTAAATGAAATAAC	CTATAAGCGTACGACTTACGGAAGTAATGTTACTTTAGGTTTCCCTGTAAATGAAAATAAC
1516	1775	1815	1831	1827		1577	· · · ·	1836	1876	1892	1888	
cad15	minnad15	eagand15	pakd15	sb33d15	consensus	ር ር	Cadi	minnad15	eagand15	pakd15	sb33d15	consensus

	1760 TGATTTTCTTTTGGTTGGAACTATAACAGCCTTAATAGAGGCTATTTCCCAACTAAAGGG		TGATTTTCT	TGATTTTTCTTTTGGTTGGAACTATAACAGCCTTAATAGAGGCTATTTCCCAACTAAAGGG		TGATTTTTCTTTTGGAACTATAACAGCCTTAATAGAGGCTATTTCCCCAACTAAGGG 9000000000000000000000000000000000	1 GTTAAAGCAAGTCTTGGTGGACGAGTTACTATTCCAGGTTCTGATAACAAATACTACAAAC					GTTAAAGCAAGTCTTGGTGGACGAGTTAC LATTCCAGGTTCTGATAACAAATACTACAAAC
	1760	2019	2059	2075	2071		1821	2080	2120	2136	2132	-
•	cad15	minnad15	eagand15	pakd15	sb33d15	consensus	cad15 1821	minnad15	eagand15	pakd15	sb33d15	consensus
												ė

AAAAGCATCTGCAGGATATGCAAATGGTTTTGGAAACAAGCGTTTTACCGTTCTATCAAACT	cad15 2004 TATACAGCGGTGGCATCGGTTCATTACGTGGTTTTTGCTTATGGTAGTATTGGACCTAACG	TATACAGCGGGTGGCATCGGTTCATTACGTTGCTTATGGTAGTATTGGACCTAACG	TATACAGCGGGTGCATCGGTTCATTACGTGGTTTTTGCTTATGGTAGTATTGGACCTAACG		$_{ m sb33d15}$ 2315 TATACAGGGGGGGGTTTTTGGTTCATTACGCGGTTTTTGCTTATGGTAGCATTGGGCCTAACG $\widetilde{\omega}$	TATACAGCGGGTGGCAT $c$ GGT $T$ CATTACG $t$ GGT $T$ TTGCT $T$ ATGGTAG $t$ A $T$ TGGACCTAA $c$ G $ ilde{N}$
	2004	2263	2303	2319	2315	
consensus	cad15	minnad15 2263	eagand15 2303	pakd15	sb33d15	consensus

CAATTTATGCCGAATATGGTAATGGTAGTGGTACTGGTACTTTTAAGAAGATAAGTTCTGA	CAATTTATGCCGAATATGGTAATGGTAGTGGTACTTTTAAGAAGATAAGTTCTGA	agtggtactggta(	CAATTTATGCCGAACATGGTAATGGTA         CTTTTAATAAGATAGGTA
2065	2324	2364	2380 (
cad15	minnad15	eagand15	pakd15

sb33d15 2376 CAATTTATcaaGgtCAaaaTAAT

**aaaTTTAATAAGATAAGTTCTGA** 

consensus		CAATTTATgccGaatAtggTAATggtagtggtactggtactTTTAAgAAGATAAGTTCTGA
cad15	2126	
minnad15 2385	2385	TGTGATTGGTGTAATGCAATCGCTACAGCTAGCGCAGAGTTAATTGTGCCAACTCCATTT
eagand15	2425	TGTGATTGGTGGTAATGCAATCGCTACAGCTAGCGCAGAGTTAATTGTGCCAACTCCATTT
pakd15 2429		TGTGATTGGTGTAATGCAATCACAACtGCAGAACTTATTGTACCAACTCCATTT
sb33d15 2422		TGTGATTGGTGGTAATGCAATCGCtACaGCtAGcGCAGAgtTaATTGTGCCAACTCCATTTN
consensus		TGTGATTGGTGGTAATGCAATCgCtACaGCtAGcGCAGAgtTaATTGTgCCAACTCCATTT

GTGAGCGATAAGAGCCAAAATACGGTCCGAACCTCCTTATTTGTTGATGCGGCAAGTGTTT GTGAGCGATAAGAGCCAAAATACGGTCCGAACCTCCTTATTTGTTGATGCGGCAAGTGTTT GTGAGCGATAAGAGCCAAAATACGGTCCGAACCIICCITATTTGTTGATGCGGCAAGTGITI'I 2187 2446 2486 cad15 minnad15 eagand15

•						•					
GTGAGTGATAAAAGCCAAAATACAGTCCGAACCTCCCTATTTGTTGATGCGGCAAGTGTTT 	GTGAGcGATAAgAGcCAAAATACgGTCCGAACCTCCŁTATTTGTTGATGCGGCAAGTGTTT	GGAATACTAA	GGAATACTAAATGGAAATCAGATAAAAATGGATTAGAGAGCGATGTATTAAAAAGATTGCC	GGAATACTAAATGGAAATCAGATAAAAATGGATTAGAGAGCGATGTATTAAAAAGATTGCC N	GGAATACTAAATGGAAATCAGATAAAAATGGATTAGAGAGCAAGGTCTTGAAAGACTTACC	GGAATACTAAATGGAAATCAGATAAAAATGGATTAGAGAGCAAtGTCTTGAAAGACTTACC	GGAATACTAAATGGAAATCAGATAAAAATGGATTAGAGAGCgAtGTaTTaAAAagaTTgCC		TGATTATGGCAAATCAAGCCGTATTCGCGCCTCTACAGGTGTCGGATTCCAATGGCAATCT	TGATTATGGCAAATCAAGCCGTATTCGCGCCTCTACAGGTGTCGGATTCCAATGGCAATCT	
2490 2483		2248	2507	2547	2551	2544		0	2309	2568	
pakd15 sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus		cadis	minnad15	

2431

cad15

eagand15	2608	eagand15 2608 TGATTATGGCAAATCAAGCCGTATTCGCGCCTCTACAGGTGTCGGATTCCAATGGCAATCT
pakd15	2612	TGATTATGGC
sb33d15	2605	
consensus		tGATTATGGCAAATCAAGCCGTAtTCGCGCCTCTACAGGTGTCGGATTCCAATGGCAATCT
cad15	cad15 2370	CCTATTGGGCCATTGGTATTCTCTTTATGCCAAACCAATTAAAAAATATGAAAATGATGATGOTGOTGOTGOTGOTGOTGOTGOTGOTGOTGOTGOTGOT
minnad15	2629	
eagand15	2669	CCTATTGGGCCATTGGTATTCTCTTTATGCCAAACCAATTAAAAAATATGAAAATGATGATGATGATG
pakd15 2673	2673	
sb33d15	2666	
consensus		CCTALTGGGCCALTGGTATTcTCTTATGCcAAACCAATTAAAAAATATGAAAATGATGATG

2553 ATTTAATTAAGGATATTTATCAAATGAAAACATCGCAAAAGTAACCGCACTTGCTTTAGG		ATTTAATTAAGGATATTTATCAAATGAAAACATCGCAAAAGTAACCGCACTTGCT			o atttaattaaggatatttatcaaatgaaaacatcgcaaaagtaaccgcacttgctttagg o	514 TATTGCACTTGCTTCAGGCTATGCTTCCGCTGAAGAAAATTGCTTTCATTAATGCAGGT		13 TATTGCACTTGCTTCAGGCTATGCTTCCGCTGAAAAAAATTGCTTTCATTAATGC ACT			
	5 2812	5 2852	15 2856	.5 2849	Ø	cad15 2614	.5 2873	5 2913	15 2917	5 2910	ហ
cad15	minnad15	eagand15	pakd15	sb33d15	consensns	cad	minnad15	eagand15	pakd15	sb33d15	consensus

თ ს ---gatcgccaagcggtagcagataaacttgatgctgaatttaaდ N atattTTTcaAcatCacccagatcgccaagcggtagcagataaacttgatgctgaatttaa TATAnTTTnCAAggCnaagg ttat-ttttcaaa-c--TATALTTTCAA **LTATA** cad15 2675 2934 2977 2973 2970 pakd15 minnad15 sb33d15 eagand15 consensus

cad15 2736 acctgtagctgagaaattagcagcaagcaaaaaagaagttgatgataaaattgctgctgct

minnad15 2954

eagand15 2985

pakd15 2990

sb33d15 2975

consensus

acctgtagctgagaaattagcagcaagcaaaaaaaagaagttgatgataaaattgctgctgct

# FIG.1F.(CONTINUED)

cad15 2797 cgtaaaaaagtagaagcaaaagttgcggctttagaaaaagatgcacctcgcttacgtcaag

minnad15 2954

eagand15 2985

pakd15 2990

sb33d15 2975

consensus

ctgatattcaaaaacgccaacaggagattaataaattaggtgcggctgaagatgctgaatt

minnad15 2954

2858

cad15

eagand15 2985

pakd15 2990

sb33d15 2975

consensus

ctgatattcaaaaacgccaacaggagattaataaattaggtgcggctgaagatgctgaatt

FIG.1F.(CONTINUED)

cad15 2919 acaaaaattaatgcaagaacaagataaaaa

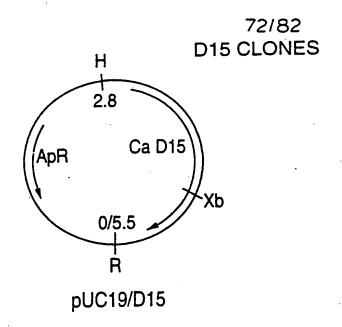
minnad15 2954 eagand15 2985

pakd15 2990

sb33d15 2975

acaaaaattaatgcaagaacaagataaaaa

consensus



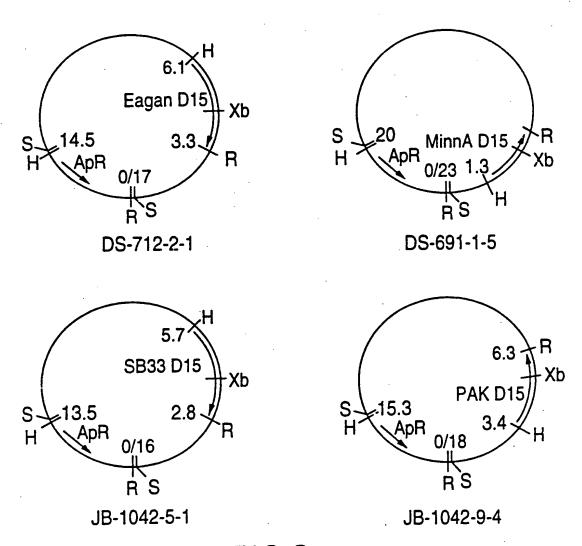


FIG.2.
SUBSTITUTE SHEET

# D15 SEQUENCE COMPARISON

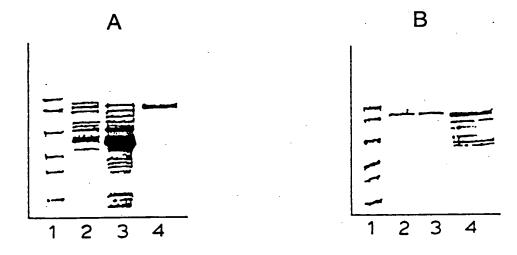
C 4	-	73/8			_	<b>A</b>
Ca Eagan MinnA SB33	PAK Ca Eagan	MirmA SB33 PAK	Ca Eagan MinnA	SB33 PAK	Ca Eaga	MinnA SB33 PAK
MRKILIASII.FGITTIVFAAPFVAKDIKVDGVQGDI.EQQIRASI.EVRAQQRVIDAUVANIVRSI.FVSGREDJVKAHQEGJVI.VVSVVAKSI.ISDVKIKGN	SVIPTEALKQNIDANGFKVGDVIJREKINEFAKSVREHYASVGRYNBITVEPIVANILPANRAEILJQINEDDKAKIASLIFKGNESVSSSTIQEQMEIQPD		SWWKLM-ENCHENDLOSIRDYYLANGYAKAQITIKIDVOLNDEKIKVNVITIDVNEGLOYDLRSARITGALGGAKSAELEFILSALHINDIFRRSDIAD		VENALKAKLGERGYGSATVINSVPDFIDDANKITATTLVVDAGRRUTVRQLRFEGNTVSADSITRQEMRQQEGTWYNSQLVELGKTRLDRRTD	NT.

									7	4/8									
ප	Eagan	MirmA	SB33	PAK	ප	Eagan	MirmA	SB33	PAK	පී	Eagan	MinnA	SB33	PAK	ය	Eagan	MinnA	SB33	PAK
PINCSNDEVIDAVYKVKERNITCSINFGIGYGIESGISYQASVKQINFLGIGAAVSLAGIKNDYGISVNLGYTEPYFTKICOSLGQNVFFENYINGKSDISS				T.II.	NYKRITYGSNVII GFPANENASYYVGI GHIYAKI SNFAL EYARAL YIQSAKFKONGIRTINDFDFSFGANYASI ARGYFPIKGVKASI GGRVITI RGSINKY					YKLSADVQGFYPLDRDHIMVVSAKASAGYANCFGNKRLPFYQTYTAGGIGSLRGFAYGSIGFNALYAEYGNGSGTGTFKKLSSDVIGGNALATASAELIV					PTPFVSDKSQNIVRISLFVDAASVANIKAKSDKNELESDVLKRLEDYCKSSRLRASIGNCFQAQSPIGPLVFSYAKPIKKYENDDVEQFQFSIGGSF**	**	**	**	**

Construction of plasmid expressing SB33 D15 ApR Xb pUC JB-1042-5-1 13.5 2.8 ApR 0/16 R/H RH ң BsrF I **ApR** D15 'ApR pUC BsrF I pRY-60-1 BsrF I Nde RH BsrF I/R R/Nde Nde-BsrF I oligos Nde ∠BsrF I ApR **ApR** D15 ApR pUC BsrF I DS-860-1-1 pT7-7 -Xb \\rangle R Nde Вġ Bg/Nde R/Rg Nde/R H / ' Nde D15 ApR DS-880-1-2 Xb FIG.4.

SUBSTITUTE SHEET

76/82
PURIFICATION OF D15 FROM A NON-TYPEABLE HAEMOPHILUS INFLUENZAE STRAIN 30



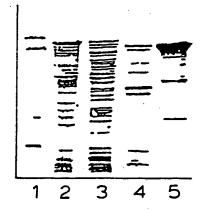
PROTEIN STAIN

**WESTERN BLOT** 

- 1. Low MW markers
- 2. Strain 30
- 3. Native D15 crude extract
- 4. D15 after anti-D15 affinity chromatography

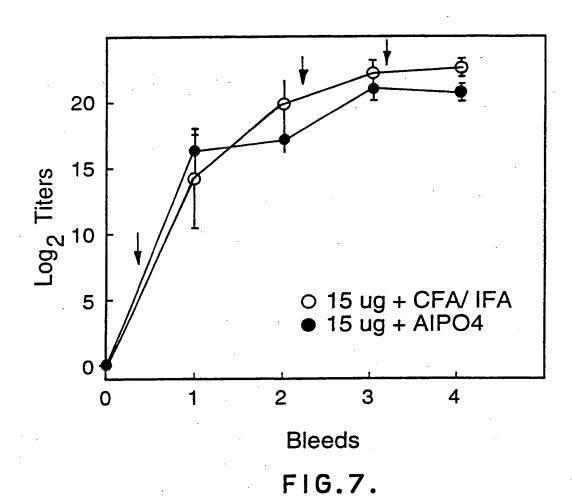
FIG.5.

# PURIFICATION OF FULL LENGTH RECOMBINANT D15



- 1. Protein M.W. Markers
- 2. Lysate of E. coli expressed rD15
- 3. Soluble protein in Tris-HC1 buffer extract
- 4. Soluble proteins in Tris/Triton X-100/ EDTA extraction buffer
- 5. rD15 inclusion bodies

FIG.6.



79/82

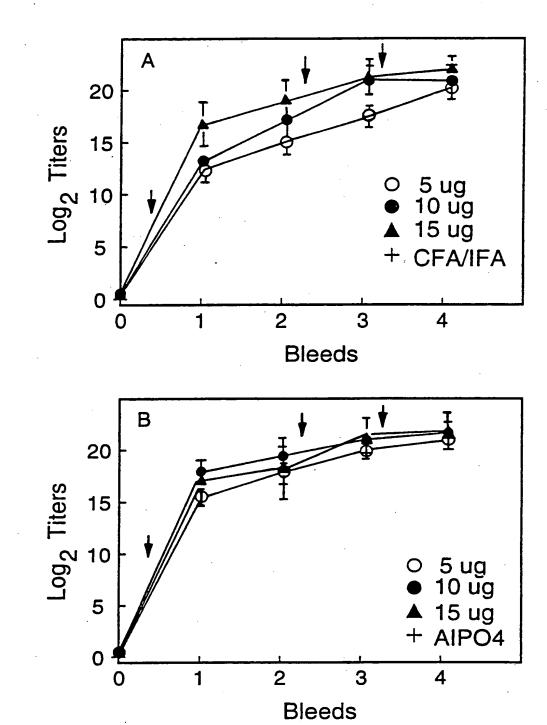
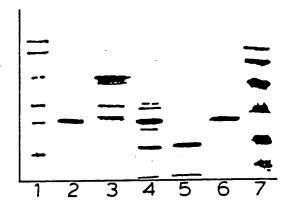


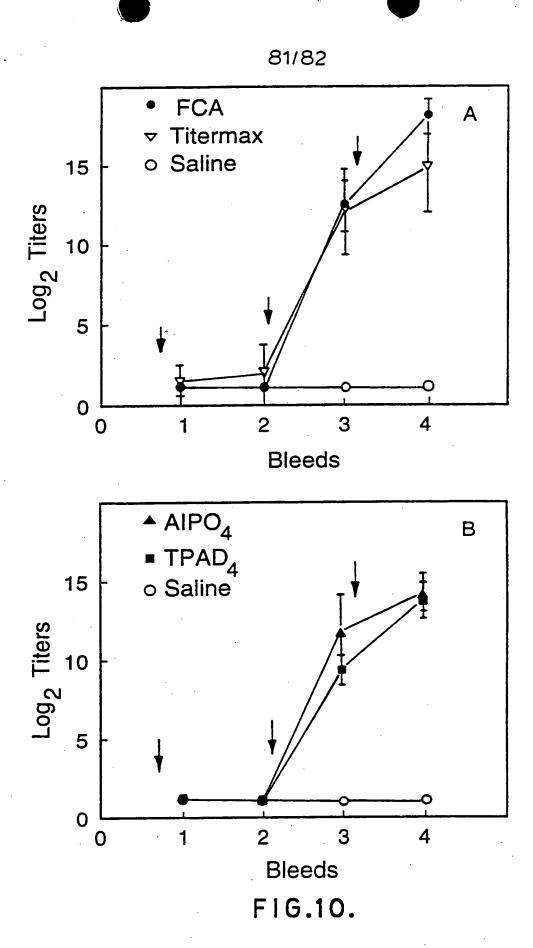
FIG.8.

PURIFICATION OF TRUNCATED D15 FROM D15-GST FUSION PROTEIN



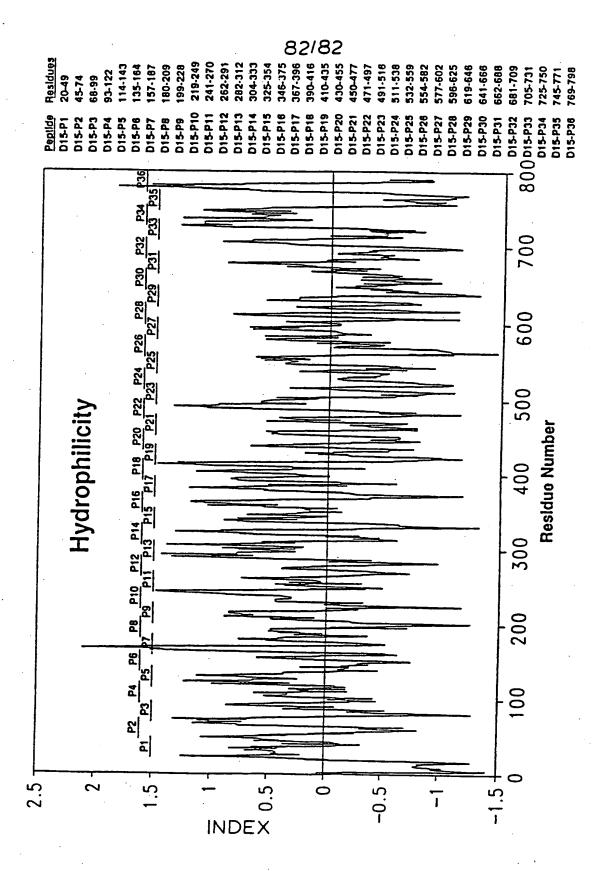
- 1. Prestain low MW markers
- 2. GST standard
- 3. GST-(D15 fragment) fusion protein
- 4. Fusion protein cleaved by thrombin
- 5. rD15 fragment
- 6. GST
- 7. Low MW markers

FIG.9.



SUBSTITUTE SHEET





# INTERNATIONAL SEARCH REPORT



tonal Application No

PCT/CA 93/00501

A. CLAS	SIFICATION OF SUBJECT MATTER					
C :	12 N 15/31,C 07 K 13/00,A 61 (C 12 N 15/31; C 12 R 1/21)	K 39/102,				
According	to International Patent Classification (IPC) or to both national ci	assification and IPC	·			
B. FIELD	S SEARCHED					
Minimum	documentation searched (classification system followed by classification s	lication symbols)				
A 6	51 K.C 07 K.C 12 N		·			
Document	mon searched other than minimum documentation to the extent t	hat such documents are included in the fields	searched			
		•				
Electronic	data hase consulted during the international search (name of data	base and, where practical, search terms used	)			
			•			
C. DOCUM	MENTS CONSIDERED TO BE RELEVANT					
Category *	Citation of document, with indication, where appropriate, of th	e tejevant passages	Relevant to claim No.			
A	INFECTION AND IMMUNITY, vol. 58, no. 4, April W.R. THOMAS et al. "I sion in Escherichia of High-Molecular-Weight Protective Surface An Found in Nontypeable Type b Haemophilus in	1,4,6, 27				
A	enzae" pages 1909-1913, the whole document.  EP, A2, 0 378 929 (CONNAUGHT LABORATOR) LIMITED) 25 July 1990 (25.07.90), claims.	1,4,6, 12,16, 20,22				
A	WO, A1, 91/06 652		1,4,6,			
Furth	er documents are listed in the continuation of box C.	Patent family membra are listed	in ameri.			
"A" documer consider the filing di filing di "L" documer which is citation "O" documer other m "P" documer later the	nt which may throw doubts on priority claim(s) or s cited to establish the publication date of another or other special reason (as specified) nt referring to an oral disclosure, use, exhibition or	T later document published after the international filling date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention.  X' document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone.  Y' document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.  '&' document member of the same patent family  Date of mailing of the international search report  25 -03- 1994				
Name and ma	tiling address of the ISA  European Patent Office, P.B. 5818 Patendaan 2  NL - 2280 HV Rijswijk  Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl.  Fax: (+ 31-70) 340-3016	Authorized officer WOLF e.h.				

### ANHANG

zum internationalen Recherchenbericht über die internationale Patentanneldung Nr.



### ANNEX

to the International Search Report to the International Patent Application No.



### NNEXE

au rapport de recherche international relatif à la demande de brevet international no

## PCT/CA 93/00501 SAE 82214

In diesem Anhang sind die Mitglieder der Patentfamilien der im obengemannten internationalen Recherchembericht cited in the above-mentioned interangeführten Patentdokumente angegeben. Diese Angaben dienen nur zur Unterrichtung und erfolgen ohne Gewähr.

This Annex lists the patent family members relating to the patent documents national search report. The Office is in no way liable for these particulars which are given merely for the purpose of information.

La présente annexe indique les memores de la famille de brevets relatifs aux documents de brevets cités dans le rapport de recherche international visée ci-dessus. Les reseignements fournis sont donnés à titre indicatif et n'engagent pas la responsibilité de l'Office.

Im Recherchenber angeführtes Patentd Patent document in search repor Document de brevet dans le rapport de n	oxument Veröffentl cited Publicat t date cité Date d	ichung Pat ion P le Mesi	glied(er) der entfamilie atent family ember(s) ore(s) de la le de brevets	Datum der Veröffentlichung Publication date Date de publication	
EP A2 378	929 25-07-	-90 GB AC CA AC CA AC EP AC JP AC GB AC	A 2006587 2006587 3 378929 2283286	22-03-89 22-06-90 23-06-90 01-08-90 20-11-90 22-02-89	
WO A1 9106	652 16-05-	-91 CA AA EP AS GB AC JP TS	500576 8924473	01-05-91 02-09-92 20-12-89 11-03-93	
EP A1 2916	673 14-09-	-88 US A DK AC DK A JP A2	6020/87	22-12-92 17-11-87 19-05-88 20-06-89	***************************************
US A 50134	664 07-05-	-91 keine	- none -	rien	